

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:35:28 ; Search time 30.6316 Seconds  
(without alignments)  
77.074 Million cell updates/sec

Title: US-09-833-196-2

Perfect score: 37

Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1069545 seqs, 262320428 residues

Total number of hits satisfying chosen parameters: 1069545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCR\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCRUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	94.6	89	12	US-10-424-599-188130
2	33	89.2	139	14	US-10-187-267A-17
3	32	86.5	376	14	US-10-156-761-8348
4	32	86.5	496	15	US-10-369-493-16984
5	31	83.8	332	12	US-10-425-114-46836
6	31	83.8	531	15	US-10-369-493-19712
7	30	81.1	495	15	US-10-369-493-11899
8	29	78.4	29	10	US-09-892-877-280
9	29	78.4	29	10	US-09-948-783-292
10	29	78.4	259	12	US-10-282-122A-54120
11	29	78.4	263	9	US-09-738-626-3727
12	29	78.4	375	9	US-09-965-529-5
13	29	78.4	375	10	US-09-969-680A-5
14	29	78.4	375	10	US-10-264-237-2289
15	29	78.4	471	15	US-10-369-493-17213

16	29	78.4	475	15	US-10-369-493-23186
17	29	78.4	4307	15	US-10-369-493-5698
18	29	78.4	4307	15	US-10-369-493-5699
19	29	78.4	4307	15	US-10-369-493-5700
20	28	75.7	131	12	US-10-424-599-27021
21	28	75.7	204	14	US-10-029-386-32508
22	28	75.7	266	9	US-09-738-626-6214
23	28	75.7	266	12	US-10-627-476-174
24	28	75.7	309	12	US-10-282-122A-60946
25	28	75.7	314	9	US-09-738-626-6809
26	28	75.7	386	12	US-10-424-599-263274
27	28	75.7	437	15	US-10-369-493-22927
28	28	75.7	455	12	US-10-424-599-194717
29	28	75.7	469	15	US-10-369-493-16544
30	28	75.7	470	14	US-10-156-761-14879
31	28	75.7	486	14	US-10-156-761-9284
32	28	75.7	510	15	US-10-369-493-20140
33	28	75.7	518	9	US-09-803-589-10
34	28	75.7	549	15	US-10-369-493-20848
35	28	75.7	551	13	US-10-105-929-16
36	28	75.7	551	14	US-10-365-227-16
37	28	75.7	648	12	US-10-425-114-72736
38	28	75.7	727	9	US-09-445-023A-12
39	28	75.7	727	14	US-10-097-597-12
40	28	75.7	727	14	US-10-097-580-12
41	28	75.7	728	9	US-09-712-363-167
42	28	75.7	785	10	US-09-733-643-2
43	28	75.7	876	14	US-10-146-473-54
44	28	75.7	894	12	US-10-425-114-56349
45	28	75.7	950	9	US-09-321-987B-4

## ALIGNMENTS

### RESULT 1

US-10-424-599-188130

; Sequence 188130, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: SVY Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 188130

; LENGTH: 89

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_1140896C.1.pap

US-10-424-599-188130

Query Match 94.6%; Score 35; DB 12; Length 89;

Best Local Similarity 87.5%; Pred. No. 1.9;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTXIRP 9

DB 78 GVTXIRP 85

↑ 100% match this pos can be s/w/p

### RESULT 2

US-10-187-267A-17

; Sequence 17, Application US/10187267A

; Publication No. US20030124679A1

; GENERAL INFORMATION:

; APPLICANT: Short, Jay M.

; APPLICANT: Paraskar, Ashish  
 ; APPLICANT: Varoglu, Mustafa  
 ; APPLICANT: Mathur, Eric J.  
 ; TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF  
 ; MAKING AND USING THEM  
 ; FILE REFERENCE: 09010-280001  
 ; CURRENT APPLICATION NUMBER: US/10/187,267A  
 ; CURRENT FILING DATE: 2003-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/301,401  
 ; PRIOR FILING DATE: 2001-06-27  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 17  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces murayamaensis ATCC 21414  
 US-10-187-267A-17

Query Match 89.2%; Score 33; DB 14; Length 139;  
 Best Local Similarity 75.0%; Pred. No. 9;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 28 GVLATIRP 35

RESULT 3  
 ; Sequence 8348, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 8348  
 ; LENGTH: 376  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces avermitilis  
 US-10-156-761-8348

Query Match 86.5%; Score 32; DB 14; Length 376;  
 Best Local Similarity 62.5%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 181 GVTEVRP 188

RESULT 4  
 ; Sequence 16984, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 16984  
 ; LENGTH: 496  
 ; TYPE: PRT  
 ; ORGANISM: Caulobacter crescentus  
 US-10-369-493-16984

Query Match 86.5%; Score 32; DB 15; Length 496;  
 Best Local Similarity 75.0%; Pred. No. 62;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 196 GLITSIRP 203

RESULT 5  
 ; US-10-425-114-46836  
 ; Sequence 46836, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 46836  
 ; LENGTH: 332  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-ZMELM017030E08\_FLI.pep  
 US-10-425-114-46836

Query Match 83.8%; Score 31; DB 12; Length 332;  
 Best Local Similarity 50.0%; Pred. No. 68;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 307 GLITSVRP 314

RESULT 6  
 ; US-10-369-493-19712  
 ; Sequence 19712, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; PLANTS WITH IMPROVED PROPERTIES  
 ; FILE REFERENCE: 38-10(52052)B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039

```
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19712
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Nitrosomonas europaea
US-10-369-493-19712

Query Match      83.8%; Score 31; DB 15; Length 531;
Best Local Similarity 75.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
      |:::|
Db      350 GAITAIRP 357

RESULT 7
US-10-369-493-11899
; Sequence 11899, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 11899
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Mesorhizobium loti
US-10-369-493-11899

Query Match      81.1%; Score 30; DB 15; Length 495;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
      |:::|
Db      195 GLITSLRP 202

RESULT 8
US-09-892-877-280
; Sequence 280, Application US/09892877
; Publication No. US20030077809A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P1
; CURRENT APPLICATION NUMBER: US/09/892,877
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/437,658
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-10
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 280
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-892-877-280

Query Match      78.4%; Score 29; DB 10; Length 29;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
      |:::|
Db      9 GILTMILRP 16

RESULT 10
US-10-282-122A-54120
; Sequence 54120, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```
QY      2 GVITXIRP 9
      |:::|
Db      9 GILTMILRP 16

RESULT 9
US-09-948-783-292
; Sequence 292, Application US/09948783
; Publication No. US20030100051A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 97 Human secreted proteins
; FILE REFERENCE: P2028P2
; CURRENT APPLICATION NUMBER: US/09/948,783
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,846
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: 09/892,877
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/437,658
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: PCT/US99/09847
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: 60/085,093
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,094
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,105
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,180
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/085,927
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,906
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,924
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,928
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,920
; PRIOR FILING DATE: 1998-05-18
; NUMBER OF SEQ ID NOS: 465
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 292
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-948-783-292

Query Match      78.4%; Score 29; DB 10; Length 29;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
      |:::|
Db      9 GILTMILRP 16

RESULT 10
US-10-282-122A-54120
; Sequence 54120, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
```

```

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07 00/280988
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3727
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3727

Query Match      78.4%; Score 29; DB 9; Length 263;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0;

QY      2  GVITXIRP 9
      |||
DB      164  GVTSIEP 171

RESULT 12
US-09-965-529-5
; Sequence 5, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 2071941CD1
US-09-965-529-5

Query Match      78.4%; Score 29; DB 9; Length 375;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0;

QY      2  GVITXIRP 9
      |||
DB      152  GILTMLRP 159

RESULT 13
US-09-969-680A-5
; Sequence 5, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS

```



```
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969,680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149,641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164,203
; PRIOR FILING DATE: 1999-11-09
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 2071941CD1
US-09-969-680A-5

Query Match      78.4%; Score 29; DB 10; Length 375;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      152 GILTMLRP 159

RESULT 14
US-10-264-237-2289
; Sequence 2289, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL31PI
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 2289
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-237-2289

Query Match      78.4%; Score 29; DB 15; Length 375;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      152 GILTMLRP 159

RESULT 15
US-10-369-493-17213
; Sequence 17213, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
```

```
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 17213
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-10-369-493-17213

Query Match      78.4%; Score 29; DB 15; Length 471;
Best Local Similarity 62.5%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      406 GILTAIRP 413

Search completed: April 1, 2004, 17:45:38
Job time : 31.6316 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 47.8947 Seconds

(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-1

Perfect score: 10

Sequence: 1 XXXXXXXXXX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_29Jan04:\*
  - 1: Geneseqp1980s:\*
  - 2: Geneseqp1990s:\*
  - 3: Geneseqp2000s:\*
  - 4: Geneseqp2001s:\*
  - 5: Geneseqp2002s:\*
  - 6: Geneseqp2003as:\*
  - 7: Geneseqp2003bs:\*
  - 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	1	2	ADD95004 Platelet
2	0	0.0	1	2	ADD94992 Platelet
3	0	0.0	1	2	ADD94993 Platelet
4	0	0.0	1	2	ADD95002 Platelet
5	0	0.0	1	2	ADD94997 Platelet
6	0	0.0	1	2	ADD94999 Platelet
7	0	0.0	1	2	ADD95003 Platelet
8	0	0.0	1	2	ADD94995 Platelet
9	0	0.0	1	2	ADD94998 Platelet
10	0	0.0	1	2	ADD95001 Platelet
11	0	0.0	1	2	ADD94990 Platelet
12	0	0.0	1	2	ADD94991 Platelet
13	0	0.0	1	2	ADD94994 Platelet
14	0	0.0	1	2	ADD94996 Platelet
15	0	0.0	1	2	ADD95000 Platelet
16	0	0.0	1	2	AAV46652 Immunogen
17	0	0.0	1	4	AAm97834 Human pep
18	0	0.0	1	4	AAm97974 Human pep
19	0	0.0	1	4	AAm97643 Human pep
20	0	0.0	1	4	AAm98447 Human pep
21	0	0.0	1	4	AAm98354 Human pep
22	0	0.0	1	4	AAm98350 Human pep
23	0	0.0	1	4	AAm53290 Human non
24	0	0.0	1	4	AAm53329 Human non
25	0	0.0	1	4	AAm53291 Human non
					AAm53328 Human non

26	0	0.0	1	4	AAm53219 Human non
27	0	0.0	1	4	AAm53218 Human non
28	0	0.0	1	4	AAb91029 Thyrotrop
29	0	0.0	1	4	AAb91739 Opicid pe
30	0	0.0	1	4	AAb92150 Polypepti
31	0	0.0	1	4	AAb91892 Apoptosis
32	0	0.0	1	4	AAb91546 Endotheli
33	0	0.0	1	4	AAb92392 Miscellan
34	0	0.0	1	4	AAb91665 Opicid pe
35	0	0.0	1	4	AAg99966 ERA bindi
36	0	0.0	1	4	AAm00011 ERA bindi
37	0	0.0	1	4	AAg99987 ERA bindi
38	0	0.0	1	4	AAg99983 ERA bindi
39	0	0.0	1	4	AAm00013 ERA bindi
40	0	0.0	1	4	AAm00016 ERA bindi
41	0	0.0	1	4	AAg99988 ERA bindi
42	0	0.0	1	4	AAm00010 ERA bindi
43	0	0.0	1	4	Abb66809 Drosophil
44	0	0.0	1	4	Abb66810 Drosophil
45	0	0.0	1	4	AAg98134 Human SNP

ALIGNMENTS

RESULT 1  
ADD95004  
ID ADD95004 standard; peptide; 1 AA.  
XX  
AC ADD95004;  
DT 29-JAN-2004 (first entry)  
XX Platelet aggregation inhibitor peptide #146.  
DE Platelet aggregation inhibitor; guanidino group; amidino group.  
KW platelet aggregation inhibitor; guanidino group; amidino group.  
XX Unidentified.  
XX Key Location/Qualifiers  
FH Modified-site I  
FT /label= OTHER  
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
XX  
XX WO9501371-Al.  
XX  
PD 12-JAN-1995.  
XX  
PF 22-JUN-1994; 94WO-JP000999.  
XX  
PR 30-JUN-1993; 93JP-00186755.  
XX  
PA (YAMA ) NIPPON STEEL CORP.  
XX  
PI Sato Y, Hayaashi Y, Katada J, Takiguchi Y;  
XX WPI; 1995-060950/08.  
XX  
DR New RGD peptide(s) useful as anti-platelet aggregation agents - contain  
XX guanidino or amidino gp. at N-terminal to increase stability.  
XX  
PS Disclosure; Page 11; 34pp; Japanese.  
XX  
CC The invention describes peptides of amino acid sequence (I) and their  
XX salts, (I) are useful as platelet aggregation inhibitors and are easily  
XX absorbed by the body. Due to the presence of the N-terminal guanidino or  
XX amidino group, (I) are provided with excellent stability so that their  
XX activity can be exhibited for an effective time after administration.  
XX Thereafter they are readily metabolised and expelled. This is the amino  
XX acid sequence of a platelet aggregation inhibitor peptide.  
XX Sequence 1 AA;  
SQ

Query Match 0.0%; Score 0; DB 2; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

RESULT 2  
ADD94992  
ID ADD94992 standard; peptide; 1 AA.  
XX  
AC ADD94992;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Platelet aggregation inhibitor peptide #134.  
XX  
KW platelet aggregation inhibitor; guanidino group; amidino group.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /label= OTHER  
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
XX  
PN WO9501371-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 22-JUN-1994; 94WO-JP000999.  
XX  
PR 30-JUN-1993; 93JP-00186755.  
XX  
PA (YAWA ) NIPPON STEEL CORP.  
XX  
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
XX  
WPI; 1995-060950/08.  
XX  
New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
guanidino or amidino gp. at N-terminal to increase stability.  
XX  
PS Disclosure; Page 10; 34pp; Japanese.  
XX  
CC The invention describes peptides of amino acid sequence (I) and their  
salts. (I) are useful as platelet aggregation inhibitors and are easily  
absorbed by the body. Due to the presence of the N-terminal guanidino or  
amidino group, (I) are provided with excellent stability so that their  
activity can be exhibited for an effective time after administration.  
Thereafter they are readily metabolised and expelled. This is the amino  
acid sequence of a platelet aggregation inhibitor peptide.  
XX  
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 2; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

RESULT 4  
ADD95002  
ID ADD95002 standard; peptide; 1 AA.  
XX  
AC ADD95002;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Platelet aggregation inhibitor peptide #144.  
XX  
KW platelet aggregation inhibitor; guanidino group; amidino group.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /label= OTHER  
FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
XX  
PN WO9501371-A1.  
XX  
PD 12-JAN-1995.  
XX  
PF 22-JUN-1994; 94WO-JP000999.  
XX  
PR 30-JUN-1993; 93JP-00186755.  
XX  
PA (YAWA ) NIPPON STEEL CORP.  
XX  
PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
XX  
WPI; 1995-060950/08.  
XX  
New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
guanidino or amidino gp. at N-terminal to increase stability.  
XX  
PS Disclosure; Page 10; 34pp; Japanese.  
XX  
CC The invention describes peptides of amino acid sequence (I) and their  
salts. (I) are useful as platelet aggregation inhibitors and are easily  
absorbed by the body. Due to the presence of the N-terminal guanidino or  
amidino group, (I) are provided with excellent stability so that their  
activity can be exhibited for an effective time after administration.  
Thereafter they are readily metabolised and expelled. This is the amino  
acid sequence of a platelet aggregation inhibitor peptide.  
XX  
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 2; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

RESULT 3  
ADD94993  
ID ADD94993 standard; peptide; 1 AA.  
XX  
AC ADD94993;

WPI: 1995-060950/08.

```

Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 7
ADD95003
ID ADD95003 standard; peptide; 1 AA.
XX
XX ADD95003;
XX
XX 29-JAN-2004 (first entry)
XX
XX Platelet aggregation inhibitor peptide #145.
XX
XX platelet aggregation inhibitor; guanidino group; amidino group.
XX
XX Unidentified.
XX
XX OS
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /label= OTHER
XX /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not
XX defined)"
XX
XX WO9501371-A1.
XX
XX
XX 12-JAN-1995.
XX
XX 22-JUN-1994; 94WO-JP000999.
XX
XX 30-JUN-1993; 93JP-00186755.
XX
XX (YAMA ) NIPPON STEEL CORP.
XX
XX
XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX WPI; 1995-060950/08.
XX
XX New RCD peptide(s) useful as anti-platelet aggregation agents - contain
XX guanidino or amidino gp. at N-terminal to increase stability.
XX
XX Disclosure; Page 11; 34pp; Japanese.
XX
XX The invention describes peptides of amino acid sequence (I) and their
XX salts. (I) are useful as platelet aggregation inhibitors and are easily
XX absorbed by the body. Due to the presence of the N-terminal guanidino or
XX amidino group, (I) are provided with excellent stability so that their
XX activity can be exhibited for an effective time after administration.
XX Thereafter they are readily metabolised and expelled. This is the amino
XX acid sequence of a platelet aggregation inhibitor peptide.
XX
XX Sequence 1 AA;
XX
XX Query Match 0.0%; Score 0; DB 2; Length 1;
XX Best Local Similarity 0.0%; Pred. No. 0;
XX Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 R 1

RESULT 8
ADD94995
ID ADD94995 standard; peptide; 1 AA.
XX
XX ADD94995;
XX
XX 29-JAN-2004 (first entry)
XX
XX
XX
XX

```

XX		platelet aggregation inhibitor peptide #137.
DE		platelet aggregation inhibitor; guanidino group; amidino group.
XX		Unidentified.
OS		
XX		
FH		Key
FT		Location/Qualifiers
PT		Modified-site 1
FT		/label= OTHER
FT		/note= "OTHER= H2NC(=NH)NH(CH2) CO. Beta-Ala (not defined)"
XX		
XX		WO9501371-Al.
XX		
PN		12-JAN-1995.
XX		
PD		22-JUN-1994; 94WO-JP000999.
XX		
PF		30-JUN-1993; 93JP-00186755.
XX		
PR		(YAWA ) NIPPON STEEL CORP.
XX		
PA		Sato Y, Hayashi Y, Katada J, Takiguchi Y;
XX		
PI		WPI; 1995-060950/08.
XX		
DR		New RDP peptide(s) useful as anti-platelet aggregation agents - contain guanidino or amidino gp. at N-terminal to increase stability.
XX		
FT		Disclosure; Page 10; 34pp; Japanese.
XX		
PS		The invention describes peptides of amino acid sequence (I) and their salts. (I) are useful as platelet aggregation inhibitors and are easily absorbed by the body. Due to the presence of the N-terminal guanidino or amidino group, (I) are provided with excellent stability so that their activity can be exhibited for an effective time after administration. Thereafter they are readily metabolised and expelled. This is the amino acid sequence of a platelet aggregation inhibitor peptide.
XX		
CC		Sequence 1 AA;
XX		
CC		Query Match 0.0%; Score 0; DB 2; Length 1;
CC		Best Local Similarity 0.0%; Pred. No. 0;
CC		Matches . 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
CC		
XX		
SQ		
QY	I X I	
DB	I R I	
RESULT 9		
ADD94998		
ID	ADD94998 standard; peptide; 1 AA.	
XX		
AC	ADD94998;	
XX		
DT	29-JAN-2004 (first entry)	
XX		
DE	Platelet aggregation inhibitor peptide #140.	
XX		
KW	platelet aggregation inhibitor; guanidino group; amidino group.	
OS		
XX	Unidentified.	
XX		
FH	Key	Location/Qualifiers
PH	Modified-site 1	/label= OTHER
FT		/note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"
FT		
XX		
PN	WO9501371-Al.	
XX		

PD 12-JAN-1995.  
 XX  
 XX 22-JUN-1994; 94WO-JP000999.  
 XX  
 PR 30-JUN-1993; 93JP-00186755.  
 XX  
 PA (YAWA ) NIPPON STEEL CORP.  
 XX  
 PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 XX  
 XX WPI; 1995-060950/08.  
 DR  
 XX  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 XX The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 XX Sequence 1 AA;  
 SQ  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 R 1  
 RESULT 10  
 ADD95001  
 ID ADD95001 standard; peptide; 1 AA.  
 XX  
 AC ADD95001;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX  
 XX Platelet aggregation inhibitor peptide #143.  
 DE  
 XX platelet aggregation inhibitor; guanidino group; amidino group.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not  
 FT defined)"  
 XX  
 PN W09501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 XX 22-JUN-1994; 94WO-JP000999.  
 PF  
 XX 30-JUN-1993; 93JP-00186755.  
 PR  
 XX (YAWA ) NIPPON STEEL CORP.  
 XX  
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 PI  
 XX WPI; 1995-060950/08.  
 DR  
 XX  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 XX The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 XX Sequence 1 AA;  
 SQ  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 R 1  
 RESULT 10  
 ADD95001  
 ID ADD95001 standard; peptide; 1 AA.  
 XX  
 AC ADD95001;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX  
 XX Platelet aggregation inhibitor peptide #143.  
 DE  
 XX platelet aggregation inhibitor; guanidino group; amidino group.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not  
 FT defined)"  
 XX  
 PN W09501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 XX 22-JUN-1994; 94WO-JP000999.  
 PF  
 XX 30-JUN-1993; 93JP-00186755.  
 PR  
 XX (YAWA ) NIPPON STEEL CORP.  
 XX  
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 PI  
 XX WPI; 1995-060950/08.  
 DR  
 XX  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX

PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 XX The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 XX Sequence 1 AA;  
 SQ  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 R 1  
 RESULT 11  
 ADD94990  
 ID ADD94990 standard; peptide; 1 AA.  
 XX  
 AC ADD94990;  
 XX  
 XX 29-JAN-2004 (first entry)  
 XX  
 XX Platelet aggregation inhibitor peptide #132.  
 DE  
 XX platelet aggregation inhibitor; guanidino group; amidino group.  
 KW  
 XX Unidentified.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not  
 FT defined)"  
 XX  
 PN W09501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 XX 22-JUN-1994; 94WO-JP000999.  
 PF  
 XX 30-JUN-1993; 93JP-00186755.  
 PR  
 XX (YAWA ) NIPPON STEEL CORP.  
 XX  
 XX Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 PI  
 XX WPI; 1995-060950/08.  
 DR  
 XX  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 XX The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 XX Sequence 1 AA;  
 SQ  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 R 1

RESULT 12  
 ADD94991  
 ID ADD94991 standard; peptide; 1 AA.  
 AC  
 AC ADD94991;  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Platelet aggregation inhibitor peptide #133.  
 XX  
 KW platelet aggregation inhibitor; guanidino group; amidino group.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
 FT  
 FT  
 FT  
 PN WO9501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 PF 22-JUN-1994; 94WO-JP000999.  
 XX  
 PR 30-JUN-1993; 93JP-00186755.  
 XX  
 PA (YAWA ) NIPPON STEEL CORP.  
 PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 XX WPI; 1995-060950/08.  
 DR  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 CC The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 R 1

RESULT 13  
 ADD94994  
 ID ADD94994 standard; peptide; 1 AA.  
 AC  
 AC ADD94994;  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Platelet aggregation inhibitor peptide #136.

Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 R 1

RESULT 14  
 ADD94996  
 ID ADD94996 standard; peptide; 1 AA.  
 AC  
 AC ADD94996;  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Platelet aggregation inhibitor peptide #138.  
 XX  
 KW platelet aggregation inhibitor; guanidino group; amidino group.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
 FT  
 FT  
 FT  
 PN WO9501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX

XX  
 KW platelet aggregation inhibitor; guanidino group; amidino group.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
 FT  
 FT  
 FT  
 PN WO9501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 PF 22-JUN-1994; 94WO-JP000999.  
 XX  
 PR 30-JUN-1993; 93JP-00186755.  
 XX  
 PA (YAWA ) NIPPON STEEL CORP.  
 PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 XX WPI; 1995-060950/08.  
 DR  
 XX New RGD peptide(s) useful as anti:platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 CC The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 R 1

RESULT 14  
 ADD94996  
 ID ADD94996 standard; peptide; 1 AA.  
 AC  
 AC ADD94996;  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Platelet aggregation inhibitor peptide #138.  
 XX  
 KW platelet aggregation inhibitor; guanidino group; amidino group.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not defined)"  
 FT  
 FT  
 FT  
 PN WO9501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX

PF 22-JUN-1994; 94WO-JP000999.  
 XX  
 PR 30-JUN-1993; 93JP-00186755.  
 XX  
 PA (YAWA ) NIPPON STEEL CORP.  
 XX  
 PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 XX  
 XX WPI; 1995-060950/08.  
 XX  
 PT New RGD peptide(s) useful as anti-platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX  
 CC The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 SQ Sequence 1 AA;  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 R 1  
 Search completed: April 1, 2004, 17:35:21  
 Job time : 48.8947 secs

Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 Db 1 R 1

## RESULT 15

ID ADD95000 standard; peptide; 1 AA.  
 XX  
 AC ADD95000;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Platelet aggregation inhibitor peptide #142.  
 XX  
 KW platelet aggregation inhibitor; guanidino group; amidino group.  
 XX  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT /note= "OTHER= H2NC(=NH)NH(CH2)CO. Beta-Ala (not  
 FT defined)"  
 XX  
 PN WO9501371-A1.  
 XX  
 PD 12-JAN-1995.  
 XX  
 PF 22-JUN-1994; 94WO-JP000999.  
 XX  
 PR 30-JUN-1993; 93JP-00186755.  
 XX  
 PA (YAWA ) NIPPON STEEL CORP.  
 XX  
 PI Sato Y, Hayashi Y, Katada J, Takiguchi Y;  
 XX  
 XX WPI; 1995-060950/08.  
 XX  
 PT New RGD peptide(s) useful as anti-platelet aggregation agents - contain  
 PT guanidino or amidino gp. at N-terminal to increase stability.  
 XX  
 PS Disclosure; Page 10; 34pp; Japanese.  
 XX

CC The invention describes peptides of amino acid sequence (I) and their  
 CC salts. (I) are useful as platelet aggregation inhibitors and are easily  
 CC absorbed by the body. Due to the presence of the N-terminal guanidino or  
 CC amidino group, (I) are provided with excellent stability so that their  
 CC activity can be exhibited for an effective time after administration.  
 CC Thereafter they are readily metabolised and expelled. This is the amino  
 CC acid sequence of a platelet aggregation inhibitor peptide.  
 XX  
 SQ Sequence 1 AA;  
 Query Match 0.0%; Score 0; DB 2; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 Db 1 R 1  
 Search completed: April 1, 2004, 17:35:21  
 Job time : 48.8947 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:30:38 ; Search time 14.2105 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-1

Perfect score: 10

Sequence: 1 XXXXXXXXXX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/aa/5A-COMB.pep.\*  
2: /cgn2\_6/prodata/2/aa/5B-COMB.pep.\*  
3: /cgn2\_6/prodata/2/aa/6A-COMB.pep.\*  
4: /cgn2\_6/prodata/2/aa/6B-COMB.pep.\*  
5: /cgn2\_6/prodata/2/aa/PCITUS-COMB.pep.\*  
6: /cgn2\_6/prodata/2/aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	1	US-07-820-154A-12
2	0	0.0	1	1	US-07-791-213D-24
3	0	0.0	1	1	US-07-791-213D-40
4	0	0.0	1	1	US-08-174-365A-57
5	0	0.0	1	1	US-07-789-913-23
6	0	0.0	1	1	US-07-789-913-25
7	0	0.0	1	1	US-08-049-794-23
8	0	0.0	1	1	US-08-049-794-25
9	0	0.0	1	1	US-08-433-037-12
10	0	0.0	1	1	US-07-869-933-16
11	0	0.0	1	1	US-08-293-150A-24
12	0	0.0	1	1	US-08-293-150A-40
13	0	0.0	1	1	US-08-496-847-23
14	0	0.0	1	1	US-08-496-847-25
15	0	0.0	1	1	US-08-742-774-23
16	0	0.0	1	2	US-08-742-774-25
17	0	0.0	1	2	US-08-675-354-23
18	0	0.0	1	2	US-08-675-354-25
19	0	0.0	1	2	US-08-097-554A-12
20	0	0.0	1	2	US-08-965-918-23
21	0	0.0	1	2	US-08-965-918-25
22	0	0.0	1	2	US-09-138-439-23
23	0	0.0	1	2	US-09-138-439-25
24	0	0.0	1	3	US-08-480-640A-12
25	0	0.0	1	3	US-08-613-400A-23
26	0	0.0	1	3	US-08-613-400A-25
27	0	0.0	1	3	US-08-613-400A-25

28	0	0.0	1	3	US-08-801-092-10	Sequence 10, Appl
29	0	0.0	1	3	US-08-801-092-17	Sequence 17, Appl
30	0	0.0	1	3	US-08-801-092-24	Sequence 24, Appl
31	0	0.0	1	3	US-08-801-092-31	Sequence 31, Appl
32	0	0.0	1	3	US-08-801-092-38	Sequence 38, Appl
33	0	0.0	1	3	US-08-801-092-45	Sequence 45, Appl
34	0	0.0	1	3	US-09-298-017-23	Sequence 23, Appl
35	0	0.0	1	3	US-09-298-017-25	Sequence 25, Appl
36	0	0.0	1	3	US-08-295-803-12	Sequence 12, Appl
37	0	0.0	1	3	US-09-392-979A-23	Sequence 23, Appl
38	0	0.0	1	3	US-09-392-979A-25	Sequence 25, Appl
39	0	0.0	1	3	US-09-103-663-16	Sequence 16, Appl
40	0	0.0	1	3	US-08-488-237A-12	Sequence 12, Appl
41	0	0.0	1	3	US-09-117-927-5	Sequence 5, Appl
42	0	0.0	1	4	US-08-375-992A-12	Sequence 12, Appl
43	0	0.0	1	4	US-09-315-113-10	Sequence 10, Appl
44	0	0.0	1	4	US-09-315-113-17	Sequence 17, Appl
45	0	0.0	1	4	US-09-315-113-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1  
US-07-820-154A-12  
; Sequence 12, Application US/07820154A  
; Patent No. 5382425  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Junker M.S., David E  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/820,154A  
; FILING DATE: 19920113  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)977-9550  
; TELEFAX: (212)664-0525  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-820-154A-12

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 V 1

RESULT 2

US-07-791-213D-24

```

; Sequence 24, Application US/07791213D
; Patent No. 5409895
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,213D
; FILING DATE: 13-NOV-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 3
US-07-791-213D-40
; Sequence 40, Application US/07791213D
; Patent No. 5409895
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 108
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States

```

```

; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/791,213D
; FILING DATE: 13-NOV-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 4
US-08-174-365A-57
; Sequence 57, Application US/08174365A
; Patent No. 5478809
; GENERAL INFORMATION:
; APPLICANT: Seiichi TANIDA et al.
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,365A
; FILING DATE: December 28, 1993
; CLASSIFICATION: S14
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 57:

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acid  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: modified site  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note = "Xaa is modified amino acid as  
OTHER INFORMATION: described in specification"  
US-08-174-365A-57

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 100.0%; Pred.No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 X 1  
Db 1 X 1

RESULT 5  
US-07-789-913-23  
Sequence 23, Application US/07789913  
Patent No. 5559095  
GENERAL INFORMATION:  
APPLICANT: Miljanich, George P.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Fox, James A.  
APPLICANT: Valentino, Karen L.  
APPLICANT: Bitner, Robert S.  
APPLICANT: Yamashiro, Donald H.  
TITLE OF INVENTION: Delayed Treatment Method of Reducing  
TITLE OF INVENTION: Ischemia-Related Neuronal Damage  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Peter Dehlinger  
STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/789,913  
FILING DATE: 19911112  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/561,766  
FILING DATE: 02-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/440,094  
FILING DATE: 22-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0005.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: peptide fragment used in the claims  
US-07-789-913-23  
Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 X 1  
Db 1 C 1  
RESULT 6  
US-07-789-913-25  
Sequence 25, Application US/07789913  
Patent No. 5559095  
GENERAL INFORMATION:  
APPLICANT: Miljanich, George P.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Fox, James A.  
APPLICANT: Valentino, Karen L.  
APPLICANT: Bitner, Robert S.  
APPLICANT: Yamashiro, Donald H.  
TITLE OF INVENTION: Delayed Treatment Method of Reducing  
TITLE OF INVENTION: Ischemia-Related Neuronal Damage  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Peter Dehlinger  
STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/789,913  
FILING DATE: 19911112  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/561,766  
FILING DATE: 02-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/440,094  
FILING DATE: 22-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0005.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: peptide fragment used in the claims  
US-07-789-913-25  
Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 X 1

Db 1 R 1

## RESULT 7

US-08-049-794-23  
 ; Sequence 23, Application US/08049794  
 ; Patent No. 5587454  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JUSTICE, ALAN  
 ; APPLICANT: SINGH, TEJINDER  
 ; APPLICANT: GOHIL, KISHOR C  
 ; APPLICANT: VALENTINO, KAREN L  
 ; APPLICANT: MILJANICH, GEORGE P  
 ; TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND  
 ; TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Law Offices of Peter Dehlinger  
 ; STREET: 350 Cambridge Avenue, Suite 300  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306

COMPUTER READABLE FORM: disk  
 MEDIUM TYPE: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/049,794  
 FILING DATE: 19930415  
 CLASSIFICATION: S14

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/814,759  
 FILING DATE: 30-DEC-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stratford, Carol A.  
 REGISTRATION NUMBER: 34,444  
 REFERENCE/DOCKET NUMBER: 5865-0009.30  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE  
 INDIVIDUAL ISOLATE: 32  
 US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 C 1

## RESULT 8

US-08-049-794-25  
 ; Sequence 25, Application US/08049794  
 ; Patent No. 5587454  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JUSTICE, ALAN  
 ; APPLICANT: SINGH, TEJINDER  
 ; APPLICANT: GOHIL, KISHOR C  
 ; APPLICANT: VALENTINO, KAREN L

APPLICANT: MILJANICH, GEORGE P  
 TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND  
 TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Law Offices of Peter Dehlinger  
 STREET: 350 Cambridge Avenue, Suite 300  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94306  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/049,794  
 FILING DATE: 19930415  
 CLASSIFICATION: S14  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/814,759  
 FILING DATE: 30-DEC-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stratford, Carol A.  
 REGISTRATION NUMBER: 34,444  
 REFERENCE/DOCKET NUMBER: 5865-0009.30  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 25:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHEICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE  
 INDIVIDUAL ISOLATE: 32  
 US-08-049-794-25

Query Match 0.0%; Score 0; DB 1; Length 1;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1

Db 1 R 1

## RESULT 9

US-08-433-037-12  
 ; Sequence 12, Application US/08433037  
 ; Patent No. 5707828  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sreekrishna, Kotikanyadan  
 ; APPLICANT: Barr, Kathryn A.  
 ; APPLICANT: Briarley, Russell A.  
 ; APPLICANT: Thill, Gregory P.  
 ; APPLICANT: Tschopp, Juerg F.  
 ; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
 ; TITLE OF INVENTION: PICHIA PASTORIS  
 ; NUMBER OF SEQUENCES: 19  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Scully, Scott, Murphy & Presser  
 ; STREET: 400 Garden City Plaza  
 ; CITY: Garden City  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 11530-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-433-037-12

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 X 1
Db 1 L 1

RESULT 10
US-08-448-606-4
; Sequence 4, Application US/08448606
; Patent No. 5721114
; GENERAL INFORMATION:
; APPLICANT: Abrahams n, Lars
; APPLICANT: Holmgren, Erik
; APPLICANT: Kaider n, Christina
; APPLICANT: Lake, Mats
; APPLICANT: Mikaelsson, sa
; APPLICANT: Sejltz, Torsten
; TITLE OF INVENTION: Expression System For Producing
; TITLE OF INVENTION: Apolipoprotein AI-M
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pollock, Vande Sande & Priddy
; STREET: 1990 M Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,606
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SE93/01061
; FILING DATE: 09-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9203753-0
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Amernick, Burton A.
; REGISTRATION NUMBER: 24,852
; REFERENCE/DOCKET NUMBER: 0151/00121
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)331-7111
; TELEFAX: (202) 293-6229
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acid
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-448-606-4

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 X 1
Db 1 Q 1

RESULT 11
US-07-869-933-16
; Sequence 16, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-869-933-16

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 X 1
Db 1 M 1

RESULT 12
US-08-293-150A-24
; Sequence 24, Application US/08293150A
```

```

; Patent No. 5792629
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,150A
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/791,213
; FILING DATE: 13-NOV-1990
; PRIOR APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 24:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-293-150A-24

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 13
US-08-293-150A-40
; Sequence 40, Application US/08293150A
; Patent No. 5792629
; GENERAL INFORMATION:
; APPLICANT: MORISHITA, Hideaki
; APPLICANT: KANAMORI, Toshinori
; APPLICANT: NOBUHARA, Masahiro
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF
; TITLE OF INVENTION: TREATING USING THE SAME
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria

```

```

; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,150A
; FILING DATE: 19-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/791,213
; FILING DATE: 13-NOV-1990
; PRIOR APPLICATION NUMBER: JP 2-306745
; FILING DATE: 13-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 029650-049
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 40:
; LENGTH: 1 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-293-150A-40

Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 X 1
Db 1 C 1

RESULT 14
US-08-496-847-23
; Sequence 23, Application US/08496847
; Patent No. 5795864
; GENERAL INFORMATION:
; APPLICANT: Amstutz, Gary A.
; APPLICANT: Bowersox, Stephen S.
; APPLICANT: Gehli, Kishorchandra
; APPLICANT: Adriaenssens, Peter I.
; APPLICANT: Kristipati, Ramasharma
; TITLE OF INVENTION: METHODS AND
; TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Denlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94306-1546
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,847
; FILING DATE: 27-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A

```

REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0009.31  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-23

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 C 1

RESULT 15  
US-08-496-847-25  
Sequence 25, Application US/08496847  
Patent No. 5795864  
GENERAL INFORMATION:  
APPLICANT: Amstutz, Gary A.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Gohil, Kishorchandra  
APPLICANT: Adriaenssens, Peter I.  
APPLICANT: Kristipati, Ramasharma  
TITLE OF INVENTION: METHODS AND  
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94306-1546  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,847  
FILING DATE: 27-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0009.31  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-25

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 X 1  
Db 1 R 1  
Search completed: April 1, 2004, 17:42:10  
Job time : 15.2105 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:35:28 ; Search time 34.0351 Seconds  
(without alignments)  
77.074 Million cell updates/sec

Title: US-09-833-196-1  
Perfect score: 10  
Sequence: 1 XXXXXXXXXX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1069545 seqs, 262320428 residues

Total number of hits satisfying chosen parameters: 1069545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	1	9	US-09-909-348-4
2	0	0.0	1	9	US-09-982-172-3
3	0	0.0	1	9	US-09-982-172-4
4	0	0.0	1	9	US-09-982-172-9
5	0	0.0	1	9	US-09-982-172-11
6	0	0.0	1	9	US-09-982-172-19
7	0	0.0	1	9	US-09-982-172-31
8	0	0.0	1	9	US-09-982-172-35
9	0	0.0	1	9	US-09-982-172-37
10	0	0.0	1	9	US-09-982-172-46
11	0	0.0	1	9	US-09-982-172-69
12	0	0.0	1	9	US-09-982-172-81
13	0	0.0	1	9	US-09-982-172-83
14	0	0.0	1	9	US-09-982-172-86
15	0	0.0	1	9	US-09-982-172-86

16	0	0.0	1	9	US-09-982-172-93	Sequence 93, Appl
17	0	0.0	1	9	US-09-982-172-95	Sequence 95, Appl
18	0	0.0	1	9	US-09-982-172-106	Sequence 106, App
19	0	0.0	1	9	US-09-982-172-112	Sequence 112, App
20	0	0.0	1	9	US-09-982-172-120	Sequence 120, App
21	0	0.0	1	9	US-09-982-172-126	Sequence 126, App
22	0	0.0	1	9	US-09-982-172-148	Sequence 148, App
23	0	0.0	1	9	US-09-982-172-149	Sequence 149, App
24	0	0.0	1	9	US-09-982-172-155	Sequence 155, App
25	0	0.0	1	9	US-09-982-172-160	Sequence 160, App
26	0	0.0	1	9	US-09-982-172-172	Sequence 172, App
27	0	0.0	1	9	US-09-982-172-173	Sequence 173, App
28	0	0.0	1	9	US-09-982-172-175	Sequence 175, App
29	0	0.0	1	9	US-09-982-172-189	Sequence 189, App
30	0	0.0	1	9	US-09-982-172-190	Sequence 190, App
31	0	0.0	1	9	US-09-982-172-191	Sequence 191, App
32	0	0.0	1	9	US-09-982-172-195	Sequence 195, App
33	0	0.0	1	9	US-09-982-172-200	Sequence 200, App
34	0	0.0	1	9	US-09-982-172-211	Sequence 211, App
35	0	0.0	1	10	US-09-809-391-395	Sequence 395, App
36	0	0.0	1	10	US-09-809-391-611	Sequence 611, App
37	0	0.0	1	10	US-09-882-171-395	Sequence 395, App
38	0	0.0	1	10	US-09-882-171-611	Sequence 611, App
39	0	0.0	1	11	US-09-833-245-184	Sequence 184, App
40	0	0.0	1	11	US-09-833-245-186	Sequence 186, App
41	0	0.0	1	11	US-09-833-245-325	Sequence 325, App
42	0	0.0	1	11	US-09-833-245-500	Sequence 500, App
43	0	0.0	1	11	US-09-833-245-744	Sequence 744, App
44	0	0.0	1	11	US-09-833-245-1045	Sequence 1045, Ap
45	0	0.0	1	11	US-09-833-245-1119	Sequence 1119, Ap

ALIGNMENTS

RESULT 1

US-09-909-348-4  
; Sequence 4, Application US/09909348  
; Patent No. US20020042373A1  
; GENERAL INFORMATION:  
; APPLICANT: Carney, Darrell H.  
; APPLICANT: Crowther, Roger S.  
; APPLICANT: Stierberg, Janet  
; APPLICANT: Bergmann, John  
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists  
; FILE REFERENCE: Of The No. US20020042373A1-Proteolytically Activated Thrombin R  
; CURRENT FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: US/09/909,348  
; PRIOR FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Peptide fragment of Thrombin  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(14)  
; OTHER INFORMATION: Xaa at position six is Glu or Gln  
; OTHER INFORMATION: Xaa at position thirteen is Phe, Met, Leu, His or  
US-09-909-348-4

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 V 1



```

RESULT 2
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERA
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982.172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 3
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERA
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982.172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 4
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERA
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283

```

```

; CURRENT APPLICATION NUMBER: US/09/982.172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 5
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBOD
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENER
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982.172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11
Query Match 0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 K 1
RESULT 6
US-09-982-172-19
; Sequence 19, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBOD
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENER
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982.172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-19

```

```
Query Match      0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 7
US-09-882-172-31
; Sequence 31, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-31

Query Match      0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 8
US-09-982-172-35
; Sequence 35, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-35

Query Match      0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 R 1

RESULT 9
```

```
US-09-982-172-37
; Sequence 37, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-37

Query Match      0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 10
US-09-982-172-46
; Sequence 46, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-46

Query Match      0.0%; Score 0; DB 9; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 X 1
Db      1 K 1

RESULT 11
US-09-982-172-69
; Sequence 69, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING
; TITLE OF INVENTION: UTILIZING EACH
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
```

; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-69

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

RESULT 12  
US-09-982-172-80  
; Sequence 80, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 80  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-80

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

RESULT 13  
US-09-982-172-81  
; Sequence 81, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 81  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-81

Query Match 0.0%; Score 0; DB 9; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

RESULT 14  
US-09-982-172-83  
; Sequence 83, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 83  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-83

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 R 1

RESULT 15  
US-09-982-172-86  
; Sequence 86, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-86

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 K 1

Search completed: April 1, 2004, 17:45:37  
Job time : 35.0351 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 10,8772 Seconds  
(without alignments)  
88,434 Million cell updates/sec

Title: US-09-833-196-1

Perfect score: 10

Sequence: 1 XXXXXXXXXX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	3	3 RHTDTC	thyroliberin - Bom
2	0	0.0	3	3 RHPGT	thyroliberin - pig
3	0	0.0	3	3 RSHST	thyroliberin - she
4	0	0.0	3	3 A92971	thyroliberin - eas
5	0	0.0	3	3 GXHU	growth-modulating
6	0	0.0	3	3 A50898	hursin - chicken
7	0	0.0	3	3 A23751	spinal cord peptid
8	0	0.0	3	3 B23751	spinal cord peptid
9	0	0.0	3	3 A33802	thyrotropin-releas
10	0	0.0	3	3 A22565	R-phycoerythrin al
11	0	0.0	3	3 PQ0010	angiotensin-conver
12	0	0.0	3	3 S13894	histidinol dehydro
13	0	0.0	3	3 A43391	TRH-like tripeptid
14	0	0.0	3	3 E37196	bradykinin-potenti
15	0	0.0	3	3 F37196	bradykinin-potenti
16	0	0.0	3	3 I50412	gene p20K protein
17	0	0.0	3	3 PT0636	T-cell receptor be
18	0	0.0	3	3 PT0578	T-cell receptor be
19	0	0.0	3	3 PT0571	T-cell receptor be
20	0	0.0	3	3 PT0622	T-cell receptor be
21	0	0.0	3	3 I78890	tyrosine protein k
22	0	0.0	3	3 S68328	blood cell protein
23	0	0.0	3	3 T13892	cytochrome-c oxida
24	0	0.0	4	1 ECKAA	antho-Ramide neur
25	0	0.0	4	2 S18401	thyroglobulin - do
26	0	0.0	4	2 A02147	phagocytosis-stimu
27	0	0.0	4	2 A32039	tyrosine-melanocy
28	0	0.0	4	2 EGNK	cardioexcitatory n
29	0	0.0	4	2 PL0140	carbon-monoxide de

30	0	0.0	4	2	PL0146	carbon-monoxide de
31	0	0.0	4	2	A37832	phenol 2-monooxyge
32	0	0.0	4	2	A48360	gamma subunit of P
33	0	0.0	4	2	I40697	biotin A - Citroba
34	0	0.0	4	2	A61300	22K superhelical D
35	0	0.0	4	2	I57745	D-mannonate hydrol
36	0	0.0	4	2	A41890	protein D - Escher
37	0	0.0	4	2	S43014	hypothetical prote
38	0	0.0	4	2	D41554	hypothetical prote
39	0	0.0	4	2	B43848	cell surface adhes
40	0	0.0	4	2	I40505	hypothetical prote
41	0	0.0	4	2	I40870	phospholipase C (E
42	0	0.0	4	2	I40804	endoglucanase F -
43	0	0.0	4	2	T46627	hypothetical prote
44	0	0.0	4	2	S53508	starvation-induced
45	0	0.0	4	2	S17255	ribosomal protein

#### ALIGNMENTS

##### RESULT 1

RHTDTC

thyroliberin - Bombina orientalis

C:Species: Bombina orientalis

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C/Accession: A90919; A01415

R:Asanuma, T.; Nakajima, T.

Chem. Pharm. Bull. 23, 3301-3303, 1975

A>Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skin.

A/Reference number: A90919; MUID:76138399; PMID:815011

A/Accession: A90919

A/Molecule type: protein

A/Residues: 1-3 <YAS>

C:Superfamily: thyroliberin precursor

C/Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

##### RESULT 2

RHPGT

thyroliberin - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C/Accession: A01415

R:Nair, R.M.G.; Barrett, J.F.; Bowers, C.Y.; Schally, A.V.

Biochemistry 9, 1103-1106, 1970

A>Title: Structure of porcine thyrotropin releasing hormone.

A/Reference number: A90560; MUID:70136150; PMID:4984938

A/Accession: A01415

A/Molecule type: protein

A/Residues: 1-3 <NAI>

R:Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.

Biochem. Biophys. Res. Commun. 37, 705-710, 1969

A>Title: The identity of chemical and hormonal properties of the thyrotropin releasing

A/Reference number: A90167; MUID:70039904; PMID:4982117

A/Contents: annotation

A/Note: biological activities and Rf values (in 17 chromatographic systems) of the synt

C:Superfamily: thyroliberin precursor

C/Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid

F1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;



B23751  
spinal cord peptide SCP-5 - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: B23751  
R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A:Reference number: A23751; PMID:85250425; PMID:4015098  
A:Accession: B23751  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <HSI>  
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 M 1

RESULT 9  
A23802  
thyrotropin-releasing hormone-like peptide - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A23802  
R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.  
J. Biol. Chem. 264, 7788-7791, 1989  
A:Title: A novel peptide, pyroglutamylglutamyproline amide, in the rabbit prostate comp  
A:Reference number: A23802; PMID:85255196; PMID:2458305  
A:Accession: A23802  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <COC>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 Q 1

RESULT 10  
A22565  
R-phycoerythrin alpha-1 chain - red alga (Gastrocloium coulteri) (fragment)  
C:Species: Gastrocloium coulteri  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; PMID:85182601; PMID:3886644  
A:Accession: A22565  
A:Molecule type: protein  
A:Residues: 1-3 <KLO>

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 2 Y 2

B23751  
spinal cord peptide SCP-5 - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: B23751  
R:Hsi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A:Reference number: A23751; PMID:85250425; PMID:4015098  
A:Accession: B23751  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <HSI>  
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 M 1

RESULT 9  
A23802  
thyrotropin-releasing hormone-like peptide - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A23802  
R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.  
J. Biol. Chem. 264, 7788-7791, 1989  
A:Title: A novel peptide, pyroglutamylglutamyproline amide, in the rabbit prostate comp  
A:Reference number: A23802; PMID:85255196; PMID:2458305  
A:Accession: A23802  
A:Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <COC>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; pyroglutamic acid  
F:1/Modified site: pyroglutamate carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 Q 1

RESULT 10  
A22565  
R-phycoerythrin alpha-1 chain - red alga (Gastrocloium coulteri) (fragment)  
C:Species: Gastrocloium coulteri  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; PMID:85182601; PMID:3886644  
A:Accession: A22565  
A:Molecule type: protein  
A:Residues: 1-3 <KLO>

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 2 Y 2

RESULT 11  
PQ0010  
angiotensin-converting enzyme inhibitor (FLP-3) - common fig  
N:Alternate names: ficus latex peptide 3  
C:Species: Ficus carica (common fig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: PQ0010  
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
Agric. Biol. Chem. 53, 2763-2767, 1989  
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
A:Reference number: PQ0008  
A:Accession: PQ0010  
A:Molecule type: protein  
A:Residues: 1-3 <MAR>  
A:Experimental source: latex  
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 L 1

RESULT 12  
S13894  
histidinol dehydrogenase (EC 1.1.1.23) - wild cabbage (fragment)  
C:Species: Brassica oleracea (wild cabbage)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: S13894  
R:Nagai, A.; Scheidegger, A.  
Arch. Biochem. Biophys. 284, 127-132, 1991  
A:Title: Purification and characterization of histidinol dehydrogenase from cabbage.  
A:Reference number: S13894; PMID:91112783; PMID:1989490  
A:Accession: S13894  
A:Molecule type: protein  
A:Residues: 1-3 <NAG>  
A:Experimental source: var. capitata  
C:Keywords: dimer; NAD; oxidoreductase

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 X 1

RESULT 13  
A43391  
TSH-like tripeptide - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A43391  
R:Lackey, D.B.  
J. Biol. Chem. 267, 17508-17511, 1992  
A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroglu  
A:Reference number: A43391; PMID:92388092; PMID:1517203  
A:Accession: A43391  
A:Molecule type: protein  
A:Residues: 1-3 <LAC>  
C:Keywords: amidated carboxyl end; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 14

E37196  
bradykinin-potentiating peptide 5 - island jararaca  
C:Species: Bothrops insularis (island jararaca)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: E37196  
R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A:Title: Primary structure and biological activity of bradykinin potentiating peptides  
A:Reference number: A37196; MUID:90351557; PMID:2386615  
A:Accession: E37196  
A>Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <CIN>  
C:Keywords: pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

RESULT 15

F37196  
bradykinin-potentiating peptide 6 - island jararaca  
C:Species: Bothrops insularis (island jararaca)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: F37196  
R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A:Title: Primary structure and biological activity of bradykinin potentiating peptides  
A:Reference number: A37196; MUID:90351557; PMID:2386615  
A:Accession: F37196  
A>Status: Preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <CIN>  
C:Keywords: pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

Search completed: April 1, 2004, 17:40:36  
Job time : 11.8772 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:18 ; Search time 6.14035 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-1  
Perfect score: 10  
Sequence: 1 XXXXXXXXXXXX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	1 GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	1 LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3	1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4	1 ACHI_ACHFU	P35904 achatina fu
5	0	0.0	4	1 DCML_PSECH	P19316 pseudomonas
6	0	0.0	4	1 DMS_PSECH	P19318 pseudomonas
7	0	0.0	4	1 BOSI_HUMAN	P02731 homo sapien
8	0	0.0	4	1 PAR3_HIRME	P42562 hirudo medi
9	0	0.0	4	1 PAR4_HIRME	P42563 hirudo medi
10	0	0.0	4	1 FEKA_ANTEL	P58705 anthopleura
11	0	0.0	4	1 PLRF_HIRME	P42561 hirudo medi
12	0	0.0	4	1 FLRN_ANTEL	P58707 anthopleura
13	0	0.0	4	1 FMRF_MACNI	P01162 macrocallis
14	0	0.0	4	1 FYRI_ANTEL	P58706 anthopleura
15	0	0.0	4	1 OCPI_OCTMI	P58648 octopus min
16	0	0.0	4	1 OCPI_OCTMI	P58649 octopus min
17	0	0.0	4	1 RM01_YEAST	P36515 saccharomyc
18	0	0.0	4	1 TUPT_HUMAN	P01858 homo sapien
19	0	0.0	5	1 ALI4_CARMA	P81817 carcinus ma
20	0	0.0	5	1 BIOA_CITFR	P13071 citrobacter
21	0	0.0	5	1 BIOB_CITFR	P12997 citrobacter
22	0	0.0	5	1 BPP7_BOTIN	P30425 bothrops in
23	0	0.0	5	1 E103_LITRU	P82099 litoria rub
24	0	0.0	5	1 E104_LITRU	P82100 litoria rub
25	0	0.0	5	1 FARP_ARTTR	P41853 artiposethi
26	0	0.0	5	1 FAP2_PARMA	P81864 pardachirus
27	0	0.0	5	1 PRCT_PERAM	P01373 periplaneta
28	0	0.0	5	1 PSK_DAUCA	P58261 daucus caro
29	0	0.0	5	1 RE11_LITRU	P82070 litoria rub
30	0	0.0	5	1 RE21_LITRU	P82071 litoria rub
31	0	0.0	5	1 RE31_LITRU	P82072 litoria rub
32	0	0.0	5	1 RE32_LITRU	P82073 litoria rub
33	0	0.0	5	1 SUGA_ACHDO	P19991 acheta dome

34 0 0.0 5 1 TPIS\_CANFA P54714 canis famil  
35 0 0.0 5 1 TRM3\_ECOLI P13973 escherichia  
36 0 0.0 5 1 UC22\_MAIZE P80628 zea mays (m  
37 0 0.0 5 1 UF01\_MOUSE P38639 mus musculu  
38 0 0.0 5 1 UX44\_CHLTR P38005 chlamydia t  
39 0 0.0 6 1 ACPH\_RABIT P25194 oryctolagus  
40 0 0.0 6 1 ASP2\_LACSN P82655 lactobacill  
41 0 0.0 6 1 CIP1\_MYTED P13736 mytilus edu  
42 0 0.0 6 1 CIP2\_MYTED P13737 mytilus edu  
43 0 0.0 6 1 E101\_LITRU P82096 litoria rub  
44 0 0.0 6 1 FARP\_MONEX P41966 moniezia ex  
45 0 0.0 6 1 LOKI\_LOOMI P41491 locusta mig

## ALIGNMENTS

RESULT 1  
GRWM\_HUMAN STANDARD; PRT; 3 AA.  
ID GRWM\_HUMAN  
AC P01157;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last annotation update)  
DE Growth-modulating peptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77162369; PubMed=858356;  
RA Schlesinger D.H., Pickart L., Thaler M.M.;  
RT "Growth-modulating serum tripeptide is Glycyl-histidyl-lysine."  
RL Experientia 33:324-325(1977).  
CC -!- MISCELLANEOUS: This serum tripeptide has been found to stimulate  
CC growth of some cell types and to inhibit other types in vitro.  
CC GO: GO:0001558; P:regulation of cell growth; NAS.  
DR GO: GO:0001558; P:regulation of cell growth; NAS.  
SQ SEQUENCE 3 AA; 340 MW; 6331E81000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 G 1

RESULT 2  
LUXE\_VIBFI STANDARD; PRT; 3 AA.  
ID LUXE\_VIBFI  
AC P24272;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Long-chain-fatty-acid-luciferin-component ligase (EC 6.2.1.19) (Acyl-  
protein synthetase) (Fragment).  
GN LUXE  
OS Vibrio fischeri.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=668;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9107226; PubMed=2254256;  
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;  
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination  
site for the lux operon."  
RL J. Bacteriol. 172:6797-6802(1990).  
CC -!- FUNCTION: ACRYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE  
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS



CC SUBSTRATE IN THE LUCIPERASE-CATALYZED REACTION.  
 CC -!- CATALYTIC ACTIVITY: ATP + an acid + protein = AMP + diphosphate +  
 CC an acyl-protein thioester.  
 CC -!- PATHWAY: Bioluminescent fatty acid reduction system; second step.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M62812; -; NOT ANNOTATED\_CDS.  
 CC LUMINESCENCE; Ligase.  
 CC NON TER  
 CC SEQUENCE 3 AA; 374 MW; 6A33303000000000 CRC64;  
 CC  
 CC Query Match 0.0%; Score 0; DB 1; Length 3;  
 CC Best Local Similarity 0.0%; Pred. No. 0;  
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 X 1  
 CC Db 1 I 1  
 CC  
 CC RESULT 3  
 CC ID THYL\_PIG STANDARD; PRT; 3 AA.  
 CC AC P01151;  
 CC DT 21-JUL-1986 (Rel. 01, Created)  
 CC DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Thyloliberin (Thyrotropin releasing hormone) (TRH) (Protirelin).  
 CC OS Sus scrofa (Pig).  
 CC OS Ovis aries (Sheep).  
 CC OS Bombina orientalis (Oriental fire-bellied toad), and  
 CC OS Notopterus viridescens (Eastern newt) (Triturus viridescens).  
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 CC OX NCBI\_TaxID=9823, 9940, 8346, 8316;  
 CC [1]  
 CC RN  
 CC RP SEQUENCE.  
 CC RC SPECIES-Pig; TISSUE=Hypothalamus;  
 CC RX MEDLINE=70136150; PubMed=4984938;  
 CC RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
 CC RT "Structure of porcine thyrotropin releasing hormone."  
 CC RL Biochemistry 9:1103-1106(1970).  
 CC [2]  
 CC RN  
 CC RP SYNTHESIS.  
 CC RC SPECIES-Pig;  
 CC RX MEDLINE=70039904; PubMed=4982117;  
 CC RA Boler J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
 CC RT "The identity of chemical and hormonal properties of the thyrotropin  
 CC RT releasing hormone and pyroglutamyl-histidyl-proline amide."  
 CC RL Biochem. Biophys. Res. Commun. 37:705-710(1969).  
 CC [3]  
 CC RN  
 CC RP SEQUENCE.  
 CC RC SPECIES=Sheep; TISSUE=Hypothalamus;  
 CC RA Desiderio D.M. Jr., Burgus R., Dunn T.F., Vale W., Guillemin R.,  
 CC RA Ward D.N.;  
 CC RT "The elucidation of the primary structure of the hypothalamic thyroid  
 CC RT stimulating hormone releasing factor of ovine origin by means of mass  
 CC RT spectrometry."  
 CC RL Org. Mass Spectrom. 5:221-228(1971).  
 CC [4]  
 CC RN  
 CC RP SYNTHESIS.  
 CC RC SPECIES=Sheep;  
 CC RX MEDLINE=70163386; PubMed=4985794;  
 CC RA Burgus R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,  
 CC RA Guillemin R.;  
 CC RT "Characterization of ovine hypothalamic hypophysiotropic

RT TSH-releasing factor.";  
 RL Nature 226:321-325(1970).  
 [5]  
 RN  
 RP SEQUENCE.  
 RC SPECIES=B.orientalis; TISSUE=Skin;  
 RX MEDLINE=76138399; PubMed=815011;  
 RA Yasuhara T., Nakajima T.;  
 RT "Letter: Occurrence of Pyr-His-Pro-NH2 in the frog skin."  
 RL Chem. Pharm. Bull. 23:3301-3303(1975).  
 [6]  
 RN  
 RP SEQUENCE.  
 RC SPECIES=N.viridescens;  
 RX MEDLINE=75035605; PubMed=4214528;  
 RA Grimm-Joergensen Y., McKelvy J.F.;  
 RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus  
 RT viridescens) brain in vitro. Isolation and characterization of  
 RT thyrotropin releasing factor."  
 RL J. Neurochem. 23:471-478(1974).  
 CC -!- FUNCTION: TRH functions as a regulator of the biosynthesis of TSH  
 CC in the anterior pituitary gland and as a neurotransmitter/  
 CC neuromodulator in the central and peripheral nervous systems.  
 CC  
 CC PIR; A90919; RHTDPO.  
 CC DR PIR; A92971; A92971.  
 CC DR PIR; A93750; RSHST.  
 CC KW Amidation; Pyrrolidone carboxylic acid.  
 CC FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 CC FT MOD\_RES 3 3 AMIDATION.  
 CC SQ SEQUENCE 3 AA; 380 MW; 7761F6E000000000 CRC64;  
 CC  
 CC Query Match 0.0%; Score 0; DB 1; Length 3;  
 CC Best Local Similarity 0.0%; Pred. No. 0;  
 CC Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC  
 CC QY 1 X 1  
 CC Db 1 Q 1  
 CC  
 CC RESULT 4  
 CC ID ACHL\_ACHFU STANDARD; PRT; 4 AA.  
 CC AC P35904;  
 CC DT 01-JUN-1994 (Rel. 29, Created)  
 CC DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 CC DE Achatin-I.  
 CC OS Achatina fulica (Giant African snail).  
 CC OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
 CC OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.  
 CC OX NCBI\_TaxID=6530;  
 CC [1]  
 CC RN  
 CC RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
 CC RC STRAIN=Ferussac; TISSUE=Ganglion;  
 CC RX MEDLINE=89273551; PubMed=2597281;  
 CC RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,  
 CC RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,  
 CC RA Novales E.T., Kanpi C.G., Takeuchi H., Nomoto K.;  
 CC RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina  
 CC RT fulica Ferussac containing a D-amino acid residue."  
 CC RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
 CC [2]  
 CC RN  
 CC RP CHARACTERIZATION.  
 CC RC STRAIN=Ferussac; TISSUE=Heart atrium;  
 CC RX MEDLINE=9124856; PubMed=1675568;  
 CC RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
 CC RA Yoshida M., Harada A., Mureoka Y., Kobayashi M.;  
 CC RT "Purification of achatin-I from the atria of the African giant snail,  
 CC RT Achatina fulica, and its possible function."  
 CC RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
 CC [3]  
 CC RN  
 CC RP X-RAY CRYSTALLOGRAPHY.  
 CC RX MEDLINE=93014529; PubMed=1399265;  
 CC RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

```

RA Iwashita T., Nomoto K.;
RT "Crystal structure and molecular conformation of achain-I
RT (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a
RT D-amino acid residue.";
RL Int. J. Pept. Protein Res. 39:258-264(1992).
CC -1- FUNCTION: Neuroexcitatory peptide; increases the impulse frequency
CC and produces a spike broadening of the identified heart excitatory
CC neuron (PON); also enhances the amplitude and frequency of the
CC heart beat. Has also an effect on several other muscles.
DR PIR; A32480; A32480.
KW Hormone; D-amino acid.
FT MOD RES 2
SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 3 A 3

RESULT 5
DCML_PSECH STANDARD; PRT; 4 AA.
ID DCML_PSECH STANDARD; PRT; 4 AA.
AC P19916;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit L) (CO-DH L) (Fragment).
GN CUTL.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=230; [1]
RN SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -1- COFACTOR: Molybdenum (molybdopterin).
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P10140; P10140.
KW Oxidoreductase; Molybdenum.
FT NON TER 4
SQ SEQUENCE 4 AA; 441 MW; 7761B876F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 M 1

RESULT 6
DCMS_PSECH STANDARD; PRT; 4 AA.
ID DCMS_PSECH STANDARD; PRT; 4 AA.
AC P19918;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO
DE dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290; [1]
RN SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -1- COFACTOR: Molybdenum (molybdopterin).
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P10140; P10140.
KW Oxidoreductase; Molybdenum.
FT NON TER 4
SQ SEQUENCE 4 AA; 441 MW; 7761B876F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 M 1

RESULT 7
ECSE_HUMAN STANDARD; PRT; 4 AA.
ID ECSE_HUMAN STANDARD; PRT; 4 AA.
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilactin peptides.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606; [1]
RN SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilactin tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -1- MISCELLANEOUS: These peptides are released from mast cells in lung
CC (and other tissues) during hypersensitivity reactions
CC (anaphylaxis). Their activities, preferentially affecting
CC eosinophils, include chemotaxis, chemotactic deactivation, release
CC of enzymes, and stimulation of the hexose monophosphate shunt.
DR GO; GO:0006935; P:chemotaxis; IDA.
DR GO; GO:0006955; P:immune response; IDA.
FT VARIANT 1
FT V -> A (IN OTHER PEPTIDE).
FT /FTID=VAR_005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 3 S 3

```

```

DE dehydrogenase subunit S) (CO-DH S) (Fragment).
GN CUTS.
OS Pseudomonas carboxydohydrogena.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae.
OX NCBI_TaxID=290; [1]
RN SEQUENCE.
RX MEDLINE=90055678; PubMed=2818128;
RA Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT "Homology and distribution of CO dehydrogenase structural genes in
RT carboxydrotrophic bacteria.";
RL Arch. Microbiol. 152:335-341(1989).
CC -1- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon
CC dioxide.
CC -1- CATALYTIC ACTIVITY: CO + H(2)O + acceptor = CO(2) + reduced
CC acceptor.
CC -1- COFACTOR: Binds 2 2Fe-2S clusters.
CC -1- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC SMALL.
DR PIR; P10146; P10146.
KW Oxidoreductase; Metal-binding; Iron-sulfur; Iron; 2Fe-2S.
FT NON TER 4
SQ SEQUENCE 4 AA; 420 MW; 6DD33DD6F0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 2 A 2

RESULT 7
ECSE_HUMAN STANDARD; PRT; 4 AA.
ID ECSE_HUMAN STANDARD; PRT; 4 AA.
AC P02731;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Eosinophilactin peptides.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606; [1]
RN SEQUENCE.
RX MEDLINE=76078412; PubMed=1060093;
RA Goetzl E.J., Austen K.F.;
RT "Purification and synthesis of eosinophilactin tetrapeptides of
RT human lung tissue: identification as eosinophil chemotactic factor of
RT anaphylaxis.";
RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).
CC -1- MISCELLANEOUS: These peptides are released from mast cells in lung
CC (and other tissues) during hypersensitivity reactions
CC (anaphylaxis). Their activities, preferentially affecting
CC eosinophils, include chemotaxis, chemotactic deactivation, release
CC of enzymes, and stimulation of the hexose monophosphate shunt.
DR GO; GO:0006935; P:chemotaxis; IDA.
DR GO; GO:0006955; P:immune response; IDA.
FT VARIANT 1
FT V -> A (IN OTHER PEPTIDE).
FT /FTID=VAR_005201.
SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred.No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 3 S 3

```

```

RESULT 8
FAR3_HIRME          STANDARD;          PRT;          4 AA.
ID FAR3_HIRME
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRPamide-like neuropeptide YLRP-amide.
OS Hirudo medicinalis (Medicinal leech)
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: Belongs to the FARP (FMRPamide related peptide)
family.
CC Neuropeptide; Amidation.
KW MOD RES 4 4 AMIDATION.
FT MOD RES 4 AA; 598 MW; 69D4073B30000000 CRC64;
SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Y 1

RESULT 9
FAR4_HIRME          STANDARD;          PRT;          4 AA.
ID FAR4_HIRME
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRPamide-like neuropeptide YMRP-amide.
OS Hirudo medicinalis (Medicinal leech)
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RN SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -!- SIMILARITY: Belongs to the FARP (FMRPamide related peptide)
family.
CC Neuropeptide; Amidation.
KW MOD RES 4 4 AMIDATION.
FT MOD RES 4 AA; 616 MW; 69D4068B30000000 CRC64;
SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 1 Y 1

RESULT 10
FFKA_ATEL          STANDARD;          PRT;          4 AA.
ID FFKA_ATEL
AC P58705;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Antho-KAamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynanthaea; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RN SEQUENCE.
RX MEDLINE=92028852; PubMed=1681803;
RA Nothacker H.-P., Rinehart K.L., Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenyllactyl-Phe-Lys-Ala-NH2 (Antho-KAamide), a
novel neuropeptide from sea anemones.";
RT novel neuropeptide from sea anemones.";
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RN FUNCTION.
RX MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
inhibitory neuropeptides, Antho-KAamide and Antho-Riamide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993)
CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle
groups. May be involved in the expansion phase of feeding
behaviour in sea anemones.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron specific.
DR PIR: JQ1273; JQ1273
KW Neuropeptide; Amidation.
FT MOD RES 1 1 L-3-PHENYLLACTYL.
FT MOD RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 512 MW; 6DD339CA00000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
Db 4 A 4

RESULT 11
FLRF_HIRME          STANDARD;          PRT;          4 AA.
ID FLRF_HIRME
AC P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRPamide
DE FMRPamide
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX NCBI_TaxID=6421, 27815;
RN [1]
RN SEQUENCE.
RX SPECIES=H.medicalinalis;
RC MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [2]
RN SEQUENCE.
RX SPECIES=H.trivolvis; TISSUE=Kidney;
RC MEDLINE=94286417; PubMed=7312428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRPamide-related peptides from the kidney of the snail, Helisoma
trivolvis.";
RL Peptides 15:31-36(1994).
CC -!- SIMILARITY: Belongs to the FARP (FMRPamide related peptide)
family.
KW Neuropeptide; Amidation.
FT MOD RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

```

```

Query Match      0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 F 1

RESULT 12
FLRN ANTEL
ID FLRN ANTEL STANDARD; PRT; 4 AA.
AC P58707;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Antho-RNamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RX MEDLINE=90319122; PubMed=1973541;
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
RT "Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
RT anemone neuropeptide containing an unusual amino-terminal blocking
RT group.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414 (1990).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron specific.
CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.
DR PIR; A35779; A35779.
KW Neuropeptide; Amidation.
FT MOD_RES 1 1 L-3-PHENYLACTYL.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 549 MW; 64540729A000000000 CRC64;

Query Match      0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 F 1

RESULT 13
FMRP MACNI
ID FMRP MACNI STANDARD; PRT; 4 AA.
AC P01162;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE FMRamide (Peak C) (Cardioexcitatory neuropeptide).
OS Macrocallista nimbosa (Sun-ray clam),
OS Nereis virens (Sandworm),
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroidea;
OC Veneroidea; Veneridae; Macrocallista.
OX NCBI_TaxID=6594, 6353, 6421, 27815;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;
RX MEDLINE=77215956; PubMed=877582;
RA Price D.A., Greenberg M.J.;
RT "Structure of a molluscan cardioexcitatory neuropeptide.";
RL Science 197:670-671 (1977).
RN [2]
RP SEQUENCE, AND CHARACTERIZATION.

```

```

RC SPECIES=M.nimbosa; TISSUE=Ganglion;
RX MEDLINE=78012038; PubMed=909875;
RA Price D.A., Greenberg M.J.;
RT "Purification and characterization of a cardioexcitatory neuropeptide
RT from the central ganglia of a bivalve mollusc.";
RL Prep. Biochem. 7:261-281 (1977).
RN [3]
RP SEQUENCE.
RC SPECIES=N.virens; PubMed=2342992;
RX MEDLINE=90259866; PubMed=909875;
RA Krajciak K.G., Price D.A.;
RT "Authentic FMRamide is present in the polychaete Nereis virens.";
RL Peptides 11:75-77 (1990).
RN [4]
RP SEQUENCE.
RC SPECIES=H.medicinalis; PubMed=1686933;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908 (1991).
RN [5]
RP SEQUENCE.
RC SPECIES=H.trivolvis; TISSUE=Kidney;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis.";
RL Peptides 15:31-36 (1994).
CC -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC activities include augmentation, induction, and regularization of
CC cardiac contraction.
CC -!- SIMILARITY: Belongs to the FMRP (FMRamide related peptide)
CC family.
DR PIR; A01426; ECNK.
DR PIR; A60418; A60418.
KW Neuropeptide; Amidation.
FT MOD_RES 4 4 AMIDATION.
SQ SEQUENCE 4 AA; 600 MW; 69D40699A000000000 CRC64;

Query Match      0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 F 1

RESULT 14
FYRI ANTEL
ID FYRI ANTEL STANDARD; PRT; 4 AA.
AC P58706;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Antho-Riamide I [Contains: Antho-Riamide II].
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=92270459; PubMed=1821096;
RA Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA Grimmelikhuijzen C.J.P.;
RT "Isolation of two novel neuropeptides from sea anemones: the unusual,
RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH2 and its
RT des-phenylacetyl fragment Tyr-Arg-Ile-NH2.";
RL Peptides 12:1165-1173 (1991).
RN [2]
RP FUNCTION.
RX MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;

```

RT "The expansion behaviour of sea anemones may be coordinated by two  
 RT inhibitory neuropeptides, Antho-Ramide and Antho-Ramide.",  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).  
 CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle  
 CC groups. May be involved in the expansion phase of feeding  
 CC behaviour in sea anemones.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Neuron specific.  
 KW Neuropeptide; Amidation.  
 FT CHAIN 1 4 ANTHO-RIAMIDE I.  
 FT CHAIN 2 4 ANTHO-RIAMIDE II.  
 FT MOD\_RES 1 1 L-3-PHENYLACTYL.  
 FT MOD\_RES 4 4 AMIDATION.  
 SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 F 1

## RESULT 15

OCPL OCTMI  
 ID\_OCPN OCTMI STANDARD; PRT; 4 AA.  
 AC P58648;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Cardioactive peptides Ocp-1/Ocp-2.  
 OS Octopus minor (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;  
 OC Octopodiformes; Octopoda; Incirrata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=89766;  
 RN [1]  
 RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
 RC TISSUE=Brain;  
 RX MEDLINE=20336815; PubMed=10876044;  
 RA Iwakoshi E., Hisada M., Minakata H.;  
 RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
 RT Octopus minor.";  
 RL Peptides 21:623-630(2000).  
 CC -!- FUNCTION: Cardioactive; has both positive chronotropic and  
 CC inotropic effects on the heart. Ocp-2 is a 1000 time less  
 CC active than Ocp-1.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- PTM: Ocp-2 has L-Phe instead of D-Phe.  
 CC -!- MASS SPECTROMETRY: MW=395.2; METHOD=WALDI.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2 2 D-PHENYLALANINE (IN OCP-1).  
 SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred.No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

Db 1 G 1

Search completed: April 1, 2004, 17:39:22  
 Job time : 8.14035 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 31.5789 Seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-1

Perfect score:

```
Sequence: 1 XXXXXXXXXXXX 10
```

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Minimum	DB seq	length:
Maximum	DB seq	length: 200000000

Post-processing: Minimum Match 0%

1000 Processing: Minimum Match 90  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL 25:\*

```
1: sp archea:*
```

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human: \*

5: sp\_invertebra

```
6: sp_mammal:*
```

7: sp\_nh\*: \*

8: sp\_organelle:

```
9: sp_page:*
```

```
10: sp_plant:*
11: sp_plant:*
```

```

11: sp_ydent: *
12: sp_ydent: *
13: sp_ydent: *

```

```
12: sp_virg:
13: en_vert_ehtr:
```

13: sp\_verleidlac  
14: sp\_unclassif

14: sp\_microb:  
15: sp\_virus:\*

15: sp\_bacteriap  
16: sp\_bacteriap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	0	0.0	2	5	P83570	sepia offic
2	0	0.0	4	5	P83568	sepia offic
3	0	0.0	4	11	Q08433	rattus sp.
4	0	0.0	2	5	P83073	bacillus ce
5	0	0.0	5	10	Q99007	hordeum vul
6	0	0.0	5	13	P83308	gallus gall
7	0	0.0	6	2	P83533	lactobacill
8	0	0.0	6	5	P83569	sepia offic
9	0	0.0	6	10	P82181	spinacia ol
10	0	0.0	6	10	P82541	spinacia ol
11	0	0.0	6	10	P82182	spinacia ol
12	0	0.0	7	2	Q8XMS3	klebsiella
13	0	0.0	7	2	P47505	escherichia
14	0	0.0	7	2	P70804	azotobacter
15	0	0.0	7	2	O50556	actinobacil
16	0	0.0	7	2	Q47477	escherichia

```

RESULT 2
P83568      PRELIMINARY;      PRT;      4 AA.
AC P83568; 2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Pheromone peptide ILME.
OS Sepia officinalis (Common cuttlefish).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Sepioidea; Sepiidae; Sepia.
OX NCBI_TaxID=6610;
RN [1]
RN [2]
SEQUENCE, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND MASS
SPECTROMETRY.
RC TISSUE=Egg;
RX PubMed=10944467;
RA Zatylny C., Gagnon J., Boucaud-Camou E., Henry J.;
RT "ILME: a waterborne pheromonal peptide released by the eggs of Sepia
officinalis.";
RL Biochem. Biophys. Res. Commun. 275:217-222(2000).
RN [2]
RN SEQUENCE.
RC TISSUE=Egg;
RX PubMed=12207899;
RA Zatylny C., Marvin L., Gagnon J., Henry J.;
RT "Fertilization in Sepia officinalis: the first mollusk sperm-
attracting peptide.";
RL Biochem. Biophys. Res. Commun. 286:1186-1193(2002).
CC -!- FUNCTION: HAS MYOTROPIC ACTIVITY TARGETING THE GENITAL TRACT.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- TISSUE SPECIFICITY: FOLLICLE, FULLY GROWN OOCYTE AND EGG(EC2).
CC -!- MASS SPECTROMETRY: MW=505.4; METHOD=MALDI.
DR GO; GO:0005186; P:pheromone activity; IEA.
KW Pheromone.
SQ SEQUENCE 4 AA; 505 MW; 6B169720300000000 CRC64;

Query Match      0.0%; Score 0; DB 5; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 I 1

RESULT 3
Q08433      PRELIMINARY;      PRT;      4 AA.
AC Q08433; 1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Bilirubin UDP-glucuronosyltransferase (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Gunn;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
DR EMBL; S38636; AAB19259.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
FT NON_TER 1 1
SQ SEQUENCE 4 AA; 473 MW; 633732C4200000000 CRC64;

Query Match      0.0%; Score 0; DB 11; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 1 I 1

RESULT 4
P83073      PRELIMINARY;      PRT;      5 AA.
AC P83073; 2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 88 kDa protein (Fragment).
OS Bacillus cereus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1396;
RN [1]
RN SEQUENCE.
RC STRAIN=NCIMB 11796;
RA Browne N., Dowds B.C.A.;
RL Submitted (JUL-2001) to Swiss-Prot.
FT NON_TER 5
SQ SEQUENCE 5 AA; 623 MW; 6B01AAA336F000000 CRC64;

Query Match      0.0%; Score 0; DB 2; Length 5;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 4 T 4

RESULT 5
Q99007      PRELIMINARY;      PRT;      5 AA.
AC Q99007; 1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Alpha amylase (Fragment).
GN AMY1 GENE.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91329704; PubMed=1831055;
RA Jacobsen J.V., Close T.J.;
RT "Control of transient expression of chimaeric genes by gibberellic
acid and abscisic acid in protoplasts prepared from mature barley
aleurone layers.";
RL Plant Mol. Biol. 16:713-721(1991).
DR EMBL; X54643; CAA38455.1; -.
FT NON_TER 5
SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F000000 CRC64;

Query Match      0.0%; Score 0; DB 10; Length 5;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1
DB 2 A 2

RESULT 6
P83308

```

```

ID P83308 PRELIMINARY; PRT; 5 AA.
AC P83308;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE FMRamide-like neuropeptide (LPLRF-amide).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND
RC TISSUE=Brain;
RX Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
RA "A novel active pentapeptide from chicken brain identified by
RT antibodies to FMRamide."
RL Nature 305:328-330(1983).
CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRAMIDE RELATED PEPTIDE)
CC FAMILY.
DR GO: GO:0007218; P:neuropeptide signaling pathway; TAS.
KW Neuropeptide; Amidation.
FT MOD RES 5 5
SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 0.0%; Score 0; DB 13; Length 5;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 1 L 1

RESULT 7
P83533
ID P83533 PRELIMINARY; PRT; 6 AA.
AC P83533;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Unknown protein from 2D-page (Fragment).
OS Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1625;
RN [1]
RP SEQUENCE.
RX STRAIN=DSM 20451;
RX PubMed=12112860;
RA Drews O., Weiss W., Reil G., Parlar H., Wait R., Goerg A.;
RT "High pressure effects step-wise altered protein expression in
RT Lactobacillus sanfranciscensis."
RL Proteomics 2:765-774(2002).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED MW OF THIS UNKNOWN
CC PROTEIN IS: 15 KDA.
FT NON TER 1 1
FT NON TER 6 6
SQ SEQUENCE 6 AA; 590 MW; 6DDB452D1AAC000 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 T 2

RESULT 8
P83569

```

```

ID P83569 PRELIMINARY; PRT; 6 AA.
AC P83569;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Sperm attracting peptide SepSA.
OS Sepia officinalis (Common cuttlefish).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Neocoleoidea;
OC Decapodiformes; Sepioidae; Sepiidae; Sepia.
OX NCBI_TaxID=6610;
RN [1]
RP SEQUENCE, FUNCTION, DEVELOPMENTAL STAGE, MASS SPECTROMETRY, AND
RC AMIDATION.
RX TISSUE=Egg;
RX PubMed=13207899;
RA Zatylyy C., Marlyn L., Gagnon J., Henry J.;
RT "Fertilization in Sepia officinalis: the first mollusk sperm-
RT attracting peptide."
RL Biochem. Biophys. Res. Commun. 296:1186-1193(2002).
CC -1- FUNCTION: ATTRACTS SPERM INCREASING THE CHANCES OF GAMETE
CC COLLISION.
CC -1- DEVELOPMENTAL STAGE: FIRST APPEARS IN THE OVARIAN FOLLICLES FURING
CC VITELLOGENESIS. ACCUMULATES IN THE OOCYTES BEFORE BEING SECRETED
CC DURING FERTILIZATION. EXPRESSION CONTINUES IN THE EMBEDDED OOCYTE.
CC ACCUMULATES IN THE EGG CAPSULE AFTER FERTILIZATION.
CC -1- MASS SPECTROMETRY: MW=596.6; METHOD=MALDI.
KW Amidation.
FT MOD RES 6 6
SQ SEQUENCE 6 AA; 597 MW; 72C8676AA0470000 CRC64;

Query Match 0.0%; Score 0; DB 5; Length 6;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1
Db 2 I 2

RESULT 9
P82181
ID P82181 PRELIMINARY; PRT; 6 AA.
AC P82181;
DT 01-JUN-2000 (TReMBLrel. 14, Created)
DT 01-JUN-2000 (TReMBLrel. 14, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=33562;
RN [1]
RP SEQUENCE.
RX STRAIN=cv. ALVARO; TISSUE=Leaf;
RX MEDLINE=20435798; PubMed=10874046;
RA Yamaguchi K., Subramanian A.R.;
RT "The plastid ribosomal proteins. Identification of all the proteins in
RT the 50 S subunit of an organelle ribosome (chloroplast).";
RL J. Biol. Chem. 275:28466-28482(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR GO: GO:0009507; C:chloroplast; IEA.
DR GO: GO:0019843; F:rRNA binding; IEA.
DR GO: GO:0003735; F:structural constituent of ribosome; IEA.
DR InterPro: IPR002363; Ribosomal_L10ceb.
DR PROSITE: PS01109; RIBOSOMAL_L10; PARTIAL.
KW Ribosomal protein; Chloroplast; rRNA-binding.
FT NON TER 6 6
SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

```



Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 A 1

RESULT 10  
 P82541  
 ID P82541 PRELIMINARY; PRT; 6 AA.  
 AC P82541;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBlrel. 25, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
 OS Spinacia oleracea (Spinach).  
 OC Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 [1]

FP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC STRAIN=CV. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435797; PubMed=10874039;  
 RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 the small subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 37:28455-28465(2000).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 DR InterPro; IPR002222; Ribosomal S19.  
 DR PROSITE; PS00323; RIBOSOMAL\_S19; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 T 1

RESULT 11  
 P82182  
 ID P82182 PRELIMINARY; PRT; 6 AA.  
 AC P82182;  
 DT 01-JUN-2000 (TREMBlrel. 14, Created)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 [1]

RP SEQUENCE.  
 RC STRAIN=CV. ALVARO; TISSUE=Leaf;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 DR InterPro; IPR002363; Ribosomal L10eub.  
 DR PROSITE; PS01109; RIBOSOMAL\_L10; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321B415B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 1 A 1

RESULT 12  
 O8KMS3  
 ID O8KMS3 PRELIMINARY; PRT; 7 AA.  
 AC O8KMS3;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Putative MerR2 protein.  
 GN MERR2  
 OS Klebsiella sp. LS13-39.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Klebsiella.  
 OX NCBI\_TaxID=143776;  
 [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=LS13-39;  
 RX MEDLINE=21804134; PubMed=11763242;  
 RA Mindlin S.Z., Kholodii G.Y., Gorlenko Z.M., Minakhina S.V.,  
 RA Minakhin L.S., Kalyaeva E.S., Kopteva A.V., Petrova M.A.,  
 RA Yurieva O.V., Nikiforov V.G.;  
 RT "Mercury resistance transposons of Gram-negative environmental  
 bacteria and their classification.";  
 RL Res. Microbiol. 152:811-822(2001).  
 DR EMBL; AJ302776; CAC82975.1;  
 SQ SEQUENCE 7 AA; 608 MW; 6DC1B5BDD87DD6F0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
 DB 2 A 2

RESULT 13  
 Q47505  
 ID Q47505 PRELIMINARY; PRT; 7 AA.  
 AC Q47505;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE MccA protein.

```

GN  MCCA.
OS  Escherichia coli.
OG  Plasmid pMccC7.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Escherichia.
OX  NCBI_TaxID=562;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=96099297; PubMed=8522520;
RA  Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;
RT  "Structure and organization of plasmid genes required to produce the
RT  translation inhibitor microcin C7.";
RL  J. Bacteriol. 177:7131-7140(1995).
DR  EMBL; X57583; CAA40808.1; -
DR  PIR; S45311; S45311
DR  GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW  Plasmid.
SQ  SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match      0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 X 1
Db  3 T 3

RESULT 14
P70804 PRELIMINARY; PRT; 7 AA.
AC P70804;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Algt protein (Fragment).
GN Algt.
OS Azotobacter vinelandii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=354;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RX MEDLINE=96427318; PubMed=8630682;
RA Rehm B.H.A., Ertesvag H., Valla S.;
RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algG) is
RT part of an alg gene cluster physically organized in a manner similar
RT to that in Pseudomonas aeruginosa.";
RL J. Bacteriol. 178:5884-5889(1996).
DR EMBL; X87973; CAA61230.1; -
DR NON_TER 1
FT SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match      0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 X 1
Db  2 T 2

RESULT 15
O50556 PRELIMINARY; PRT; 7 AA.
ID O50556
AC O50556;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GlyA (Fragment).
GN GlyA.
OS Actinobacillus actinomycetemcomitans (Haemophilus

```

```

OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33384;
RC MEDLINE=96355846; PubMed=8751884;
RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,
RA Kraig E.;
RT "cis Elements and trans factors are both important in strain-specific
RT regulation of the leukotoxin gene in Actinobacillus
RT actinomycetemcomitans.";
RL Infect. Immun. 64:3451-3460(1996).
DR EMBL; U51862; AAB88721.1; -
DR NON_TER 1
FT SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match      0.0%; Score 0; DB 2; Length 7;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy  1 X 1
Db  7 A 7

Search completed: April 1, 2004, 17:38:35
Job time : 33.5789 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 43.1053 Seconds  
(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-2

Perfect score: 37

Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_29Jan04:\*
  - 1: Geneseq1980s.\*
  - 2: Geneseq1990s.\*
  - 3: Geneseq2000s.\*
  - 4: Geneseq2001s.\*
  - 5: Geneseq2002s.\*
  - 6: Geneseq2003s.\*
  - 7: Geneseq2003bs.\*
  - 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	94.6	9	3 AAY67983	Aay67983 Antiangio
2	35	94.6	9	4 AAU15043	Aau15043 Antiangio
3	35	94.6	9	4 AAU15027	Aau15027 Antiangio
4	35	94.6	9	6 AAO26633	Aao26633 Anti-angi
5	35	94.6	10	4 AAU14993	Aau14993 Antiangio
6	34	91.9	8	7 ADD31274	Add31274 Angiogene
7	34	91.9	8	7 ADD31264	Add31264 Angiogene
8	34	91.9	8	7 ADD31266	Add31266 Angiogene
9	34	91.9	8	7 ADD31268	Add31268 Angiogene
10	34	91.9	8	7 ADD31273	Add31273 Angiogene
11	34	91.9	8	7 ADD31269	Add31269 Angiogene
12	34	91.9	8	7 ADD31267	Add31267 Angiogene
13	34	91.9	8	7 ADD31282	Add31282 Angiogene
14	34	91.9	8	7 ADE83640	Ade83640 Antiangio
15	34	91.9	8	7 ADE83694	Ade83694 Antiangio
16	34	91.9	8	7 ADE83632	Ade83632 Antiangio
17	34	91.9	8	7 ADE83696	Ade83696 Antiangio
18	34	91.9	8	7 ADE83628	Ade83628 Antiangio
19	34	91.9	8	7 ADE83641	Ade83641 Antiangio
20	34	91.9	8	7 ADE83631	Ade83631 Antiangio
21	34	91.9	9	3 AAY67973	Aay67973 Antiangio
22	34	91.9	9	3 AAY67978	Aay67978 Antiangio
23	34	91.9	9	3 AAY67992	Aay67992 Antiangio
24	34	91.9	9	3 AAY67975	Aay67975 Antiangio
25	34	91.9	9	3 AAY67986	Aay67986 Antiangio

26	34	91.9	9	3 AAY67996	Aay67996 Antiangio
27	34	91.9	9	3 AAY67997	Aay67997 Antiangio
28	34	91.9	9	3 AAY67995	Aay67995 Antiangio
29	34	91.9	9	3 AAY67985	Aay67985 Antiangio
30	34	91.9	9	3 AAY67976	Aay67976 Antiangio
31	34	91.9	9	3 AAY67989	Aay67989 Antiangio
32	34	91.9	9	3 AAY67974	Aay67974 Antiangio
33	34	91.9	9	3 AAY67981	Aay67981 Antiangio
34	34	91.9	9	3 AAY67982	Aay67982 Antiangio
35	34	91.9	9	3 AAY67988	Aay67988 Antiangio
36	34	91.9	9	3 AAY67984	Aay67984 Antiangio
37	34	91.9	9	3 AAY67977	Aay67977 Antiangio
38	34	91.9	9	3 AAY67994	Aay67994 Antiangio
39	34	91.9	9	3 AAY67972	Aay67972 Antiangio
40	34	91.9	9	3 AAY67991	Aay67991 Antiangio
41	34	91.9	9	3 AAY67993	Aay67993 Antiangio
42	34	91.9	9	3 AAY67990	Aay67990 Antiangio
43	34	91.9	9	3 AAY67999	Aay67999 Antiangio
44	34	91.9	9	3 AAY67987	Aay67987 Antiangio
45	34	91.9	9	3 AAY67987	Aay67987 Antiangio

ALIGNMENTS

RESULT 1  
AAY67983  
ID AAY67983 standard; peptide; 9 AA.  
XX AAY67983;  
AC AAY67983;  
XX 11-APR-2000 (first entry)  
XX Antiangiogenic peptide #12.  
XX  
KW Antiangiogenic; antithrombotic; cytostatic; antipsoriatic; arthritis;  
KW angiogenesis; inhibitor; cancer; psoriasis; infection; ocular disease;  
KW macular degeneration; diabetic retinopathy; tumour metastasis;  
KW autoimmune disease; neovascularisation; Crohn's disease; birth control;  
KW cat scratch disease.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1  
FT /label= MeGly  
FT /note= "Sarcosine (methylglycine); acetylated"  
FT Misc-difference 4  
FT /note= "D form residue"  
FT Modified-site 9  
FT /note= "Pro is modified to ProNHCH2CH3"  
XX WO9961476-A1.  
XX  
XX 02-DEC-1999.  
XX  
XX 21-MAY-1999; 99WO-US011448.  
XX  
XX 22-MAY-1998; 98US-00083745.  
XX 16-FEB-1999; 99US-00250574.  
XX 26-MAR-1999; 99US-00277466.  
XX (ABBO ) ABBOTT LAB.  
XX  
XX Henkin J, Haviv F, Bradley MF, Kalvin DM, Schneider AJ;  
XX WPI; 2000-072606/06.  
XX  
XX New anti-angiogenic peptides, used for treating e.g. cancer, arthritis,  
XX psoriasis, or angiogenesis of the eye associated with infection or  
XX surgical intervention, macular degeneration and diabetic retinopathy.  
XX  
XX Claim 12; Page 75; 223pp; English.

XX AAU67972 to AAU67999 represent novel anti-angiogenic peptides. Peptide  
CC from the present invention have the formula: A0-A1-A2-A3-A4-A5-A6-A7-A8-  
CC A9-A10 where A0 is selected from hydrogen or an acyl group; A10 is a  
CC hydroxyl group or an amino acid amide; and A1 to A9 are amino acyl  
CC residues. The peptides are used for anti-angiogenesis therapy in  
CC patients. They are used to treat e.g. cancer, arthritis, psoriasis,  
CC angiogenesis of the eye associated with infection or surgical  
CC intervention, macular degeneration and diabetic retinopathy. They can  
CC prevent tumour metastases. Further uses include treatment and prophylaxis  
CC of autoimmune diseases, various ocular diseases e.g. diabetic  
CC retinopathy, and other abnormal neovascularisation conditions of the eye,  
CC skin diseases e.g. psoriasis, diseases characterised by excessive or  
CC abnormal stimulation of endothelial cells, e.g. Crohn's disease. They can  
CC also be used as a birth control agent, inhibiting ovulation and placental  
CC establishment, and to treat diseases that have angiogenesis as a  
CC pathological consequence e.g. cat scratch disease and are useful to  
CC reduce bleeding by administration prior to surgery, especially for the  
CC treatment of tumours. They can also be used to isolate a receptor from an  
CC endothelial cell  
XX  
SQ Sequence 9 AA;

Query Match 94.6%; Score 35; DB 3; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| | | | |  
DB 2 GVITAIRP 9

RESULT 2  
AAU15043  
ID AAU15043 standard; peptide; 9 AA.

XX AAU15043;  
XX  
XX  
DT 04-DEC-2001 (first entry)  
XX  
DE Antiangiogenic peptide #70 useful for inhibiting angiogenesis.  
XX  
XX Antangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
XX arthritis; skin disease; ocular disease; diabetic retinopathy;  
XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
XX cycostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
FH Modified-site 1  
FT /label= MeGly  
FT /note= "N-methyl-glycine (sarcosine), additionally  
FT modified by N-terminal acetyl"  
FT Misc-difference 4  
FT /note= "D-form residue"  
FT Modified-site 6  
FT /note= "N-methyl serine"  
FT Modified-site 9  
FT /note= "Modified by NH-ethyl"

XX WO200138397-A1.  
XX  
XX 31-MAY-2001.  
XX  
XX 22-NOV-2000; 2000WO-US032105.  
XX  
XX 22-NOV-1999; 99US-00447099.  
XX 31-OCT-2000; 2000US-00702649.  
XX  
XX (ABBO ) ABBOTT LAB.  
XX  
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;

XX WPI; 2001-521804/57.  
XX  
XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
XX diabetic retinopathy.  
XX  
XX Claim 39; Page 99; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides  
CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
CC antiangiogenic peptides are useful for isolating a receptor from an  
CC endothelial cell. The peptides of the invention are also useful for  
CC treating cancer, arthritis, psoriasis and other skin diseases,  
CC angiogenesis of the eye associated with infection or surgical  
CC intervention and other ocular diseases, cat scratch disease, ulcers,  
CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
CC haemangiomas and capillary action within atherosclerotic plaques,  
CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
CC Osler-Webber syndrome, myocardial angiogenesis, plaque  
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
CC wound granulation, excessive or abnormal stimulation of endothelial  
CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
CC peptides are also useful as birth control agents by inhibiting ovulation  
CC and to reduce bleeding by administration before surgery. The peptides of  
CC the invention exhibit improved metabolic stability, improved  
CC pharmacokinetics, increased water solubility, and improved oral  
CC availability. The present sequence represents antiangiogenic peptide #70  
XX  
XX Sequence 9 AA;

Query Match 94.6%; Score 35; DB 4; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| | | | |  
DB 2 GVITSIRP 9

RESULT 3  
AAU15027  
ID AAU15027 standard; peptide; 9 AA.

XX AAU15027;  
XX  
XX 04-DEC-2001 (first entry)  
XX  
XX Antangiogenic peptide #54 useful for inhibiting angiogenesis.  
XX  
XX Antangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
XX arthritis; skin disease; ocular disease; diabetic retinopathy;  
XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
XX cycostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
FH Modified-site 1  
FT /label= MeGly  
FT /note= "N-methyl-glycine (sarcosine), additionally  
FT modified by N-terminal acetyl"  
FT Misc-difference 4  
FT /note= "D-form residue"  
FT Modified-site 6  
FT /note= "N-Methyl alanine"  
FT Modified-site 9  
FT /note= "Modified by NH-ethyl"

XX WO200138397-A1.  
XX  
XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032105.  
 XX 22-NOV-1999; 93US-00447099.  
 PR 31-OCT-2000; 2000US-00702649.  
 XX (ABBO ) ABBOTT LAB.  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 PT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 PT diabetic retinopathy.  
 XX Claim 39; Page 89; 95pp; English.  
 XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14374-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #54  
 XX Sequence 9 AA;  
 SQ  
 Query Match 94.6%; Score 35; DB 4; Length 9;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GVITXIRP 9  
 Db 2 GVITAIRP 9  
 RESULT 4  
 AAO25633  
 ID AAO25633 standard; peptide; 9 AA.  
 XX AC AAO25633;  
 XX 28-MAR-2003 (first entry)  
 XX Anti-angiogenic peptide #16.  
 XX Cystostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;  
 KW dermatological; immunosuppressive; cardiac; vulvar; antiulcer;  
 KW antiarteriosclerotic; angiogenesis inhibitor; cancer; arthritis;  
 KW psoriasis; angiogenesis; eye; infection; surgical intervention;  
 KW macular degeneration; diabetic retinopathy; autoimmune disease;  
 KW ocular disease; skin disease; blood vessel disease; telangiectasia;  
 KW Osler Weber Syndrome; myocardial angiogenesis; haemophilic joint;  
 KW plaque neovascularisation; angiofibroma; wound granulation; scleroderma;  
 KW atherosclerosis; intestinal adhesion; Crohn's disease; hypertrophic scar;  
 KW birth control agent; cat scratch disease; ulcer; angiogenic.  
 XX Synthetic.  
 OS  
 XX

XX Key Location/Qualifiers  
 FT Modified-site 1 /label= MeGly  
 FT Misc-difference 4 /note= "This sarcosine residue is modified by N-Ac"  
 FT Modified-site 5 /note= "This is a D-form residue"  
 FT Modified-site 9 /note= "Residue is modified to become alloThr"  
 FT /note= "Residue is modified by NHCH2CH3"  
 XX WO20028306S-A2.  
 XX 24-OCT-2002.  
 XX 10-APR-2002; 2002WO-US011027.  
 XX 11-APR-2001; 2001US-00832733.  
 XX (ABBO ) ABBOTT LAB.  
 XX Henkin J, Haviv F, Bradley MF, Douglas XM, Schneider AJ;  
 XX WPI; 2003-111805/10.  
 XX New peptides are angiogenesis inhibitors used for treating e.g. cancer,  
 PT arthritis and psoriasis.  
 XX Claim 12; Page 31; 33pp; English.  
 XX The invention relates to novel peptides for use as angiogenesis  
 CC inhibitors. Used as angiogenesis inhibitors used for treating cancer,  
 CC arthritis, psoriasis, angiogenesis of the eye associated with infection  
 CC or surgical intervention, macular degeneration and diabetic retinopathy.  
 CC The novel peptides are also used for treating autoimmune diseases, ocular  
 CC diseases, skin diseases, plaque neovascularisation, telangiectasia,  
 CC myocardial angiogenesis, angiofibroma, wound granulation, diseases  
 CC characterised by excessive or abnormal stimulation of endothelial cells  
 CC including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars. The novel peptides are also used as  
 CC birth control agents and for treating cat scratch disease and ulcers.  
 CC This sequence represents one of the angiogenic peptides of the invention  
 XX Sequence 9 AA;  
 SQ  
 Query Match 94.6%; Score 35; DB 6; Length 9;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GVITXIRP 9  
 Db 2 GVITAIRP 9  
 RESULT 5  
 AAU14993  
 ID AAU14993 standard; peptide; 10 AA.  
 XX AC AAU14993;  
 XX 04-DEC-2001 (first entry)  
 XX Antiangiogenic peptide #20 useful for inhibiting angiogenesis.  
 XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cystostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers

FT Modified-site 1 /label= Megly  
FT /note= "N-methyl-glycine (sarcosine), additionally  
FT modified by N-terminal acetyl"  
FT Misc-difference 4 /note= "D-form residue"  
FT Modified-site 6 /note= "N-Methyl serine"  
FT Modified-site 10 /note= "D-form residue, C-terminal amide"  
XX W0200138397-A1.  
XX 31-MAY-2001.  
XX 22-NOV-2000; 2000WO-US032105.  
XX 22-NOV-1999; 99US-00447099.  
XX 31-OCT-2000; 2000US-00702649.  
XX (ABBO ) ABBOTT LAB.  
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
XX WPI; 2001-521804/57.  
XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
XX diabetic retinopathy.  
XX Claim 38; Page 88; 95pp; English.  
XX The present invention relates to novel synthetic antiangiogenic peptides  
XX (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
XX antiangiogenic peptides are useful for isolating a receptor from an  
XX endothelial cell. The peptides of the invention are also useful for  
XX treating cancer, arthritis, psoriasis and other skin diseases.  
XX angiogenesis of the eye associated with infection or surgical  
XX intervention and other ocular diseases, cat scratch disease, ulcers,  
XX macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
XX haemangiomas and capillary action within atherosclerotic plaques,  
XX autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
XX Osler-Weber syndrome, myocardial angiogenesis, plaque  
XX neovascularization, telangiectasia, haemophilic joints, angiodioma,  
XX wound granulation, excessive or abnormal stimulation of endothelial  
XX cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
XX scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
XX peptides are also useful as birth control agents by inhibiting ovulation  
XX and to reduce bleeding by administration before surgery. The peptides of  
XX the invention exhibit improved metabolic stability, improved oral  
XX pharmacokinetics, increased water solubility, and improved oral  
XX availability. The present sequence represents antiangiogenic peptide #20  
XX  
SQ Sequence 10 AA;

Query Match 94.6%; Score 35; DB 4; Length 10;  
Best Local Similarity 87.5%; Pred. No. 0.099; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| | | | |  
DB 2 GVITSIRP 9

RESULT 6  
ADD31274  
ID ADD31274 standard; peptide; 8 AA.

XX ADD31274;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Angiogenesis inhibiting peptide #49.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX Synthetic.  
XX Key Location/Qualifiers  
FT Modified-site 1 /label= OTHER  
FT /note= "OTHER= N-acetyl"  
FT Misc-difference 3 /note= "D-form residue"  
FT Modified-site 5 /label= OTHER  
FT /note= "OTHER= Nme Norvalyl (Nva)"  
FT Modified-site 8 /note= "NHCH2CH3"  
XX US2003109455-A1.  
XX 12-JUN-2003.  
XX 30-OCT-2002; 2002US-00283550.  
XX 31-OCT-2001; 2001US-0335017P.  
XX (HAVI/) HAVIV F.  
XX (BRAD/) BRADLEY M F.  
XX Haviv F, Bradley MF;  
XX WPI; 2003-843101/78.  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX Claim 11; Page 25; 26pp; English.  
XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
XX has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
XX antarthritic activities. (I) or a salt of it, is used in a  
XX pharmaceutical composition for inhibiting angiogenesis or for treating  
XX cancer in a mammal. (I) Can be also used for the treatment or prevention  
XX of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
XX and degenerative arthritis. This is the amino acid sequence of an  
XX angiogenesis inhibiting peptide of the invention.

SQ Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| | | | |  
DB 1 GVITXIRP 8

RESULT 7  
ADD31264  
ID ADD31264 standard; peptide; 8 AA.

XX ADD31264;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX Angiogenesis inhibiting peptide #39.  
XX  
XX heptapeptide compound; octapeptide compound; nonapeptide compound;

KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= OTHER

FT Misc-difference 3 /note= "OTHER= N-acetyl"

FT Modified-site 5 /note= "D-form residue"

FT Modified-site 5 /label= OTHER

FT Modified-site 8 /note= "OTHER= Norvalyl (Nva)"

FT Modified-site 8 /note= "NHCH2CH3"

XX US2003109455-A1.

XX PD 12-JUN-2003.

XX PF 30-OCT-2002; 2002US-00283550.

XX PR 31-OCT-2001; 2001US-0335017P.

XX PA (HAVI/) HAVIV F.

XX PA (BRAD/) BRADLEY M F.

XX PI Haviv F, Bradley MF;

XX PI WPI; 2003-843101/78.

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.  
 XX Claim 11; Page 24; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVITXIRP 9

Db 1 GVITXIRP 8

RESULT 8

ADD31266

ID ADD31266 standard; peptide; 8 AA.

XX AC ADD31266;

XX DT 15-JAN-2004 (first entry)

XX DE Angiogenesis inhibiting peptide #41.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;

KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= OTHER

FT Misc-difference 3 /note= "OTHER= N-acetyl"

FT Modified-site 4 /note= "D-form residue"

FT Modified-site 5 /label= OTHER

FT Modified-site 5 /note= "OTHER= allothr (not defined)"

FT Modified-site 8 /label= OTHER

FT Modified-site 8 /note= "OTHER= Norvalyl (Nva)"

FT Modified-site 8 /note= "NHCH2CH3"

XX US2003109455-A1.

XX PD 12-JUN-2003.

XX PF 30-OCT-2002; 2002US-00283550.

XX PR 31-OCT-2001; 2001US-0335017P.

XX PA (HAVI/) HAVIV F.

XX PA (BRAD/) BRADLEY M F.

XX PI Haviv F, Bradley MF;

XX PI WPI; 2003-843101/78.

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.  
 XX Claim 11; Page 24; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVITXIRP 9

Db 1 GVITXIRP 8

RESULT 9

ADD31268

ID ADD31268 standard; peptide; 8 AA.

XX AC ADD31268;

XX DT 15-JAN-2004 (first entry)

XX DE Angiogenesis inhibiting peptide #43.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;

KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

XX  
 OS Synthetic.

XX Key Location/Qualifiers

XX Modified-site 1 /label= OTHER

FT /note= "OTHER= N-acetyl"

FT Misc-difference 3 /note= "D-form residue"

FT Modified-site 5 /label= OTHER

FT /note= "OTHER= Norvalyl (Nva)"

FT Misc-difference 6 /note= "D-form residue"

FT Modified-site 8 /note= "NHCH2CH3"

FT

FT

FT

FT

FT

XX US2003109455-A1.

PN 12-JUN-2003.

XX

XX 30-OCT-2002; 2002US-00283550.

XX 31-OCT-2001; 2001US-0335017P.

XX (HAVI/) HAVIV F.

PA (BRAD/) BRADLEY M F.

XX

XX Haviv F, Bradley MF;

PI WPI; 2003-843101/78.

XX

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting

PT angiogenesis, treating cancer in mammal, or for prevention of other

PT diseases such as autoimmune diseases.

XX

PS Claim 11; Page 24; 26pp; English.

XX

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)

CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and

CC antiarthritic activities. (I) or a salt of it, is used in a

CC pharmaceutical composition for inhibiting angiogenesis or for treating

CC cancer in a mammal. (I) Can be also used for the treatment or prevention

CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune

CC and degenerative arthritis. This is the amino acid sequence of an

CC angiogenesis inhibiting peptide of the invention.

XX

SQ Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;

Best Local Similarity 100.0%; Pred. NC. 1.4e+06;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVITXIRP 9

Db 1 GVITXIRP 8

RESULT 10

ADD31273

ID ADD31273 standard; peptide; 8 AA.

XX

XX ADD31273;

XX

XX 15-JAN-2004 (first entry)

XX

XX Angiogenesis inhibiting peptide #48.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;

KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;

XX

KW

KW

KW

XX

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

XX  
 OS Synthetic.

XX Key Location/Qualifiers

XX Modified-site 1 /label= OTHER

FT /note= "OTHER= N-acetyl"

FT Modified-site 2 /label= OTHER

FT /note= "OTHER= NMe"

FT Misc-difference 3 /note= "D-form residue"

FT Modified-site 5 /label= OTHER

FT /note= "OTHER= Norvalyl (Nva)"

FT Modified-site 8 /note= "NHCH2CH3"

FT

FT

FT

FT

XX US2003109455-A1.

PN 12-JUN-2003.

XX

XX 30-OCT-2002; 2002US-00283550.

XX 31-OCT-2001; 2001US-0335017P.

XX (HAVI/) HAVIV F.

PA (BRAD/) BRADLEY M F.

XX

XX Haviv F, Bradley MF;

PI WPI; 2003-843101/78.

XX

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting

PT angiogenesis, treating cancer in mammal, or for prevention of other

PT diseases such as autoimmune diseases.

XX

PS Claim 11; Page 25; 26pp; English.

XX

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)

CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and

CC antiarthritic activities. (I) or a salt of it, is used in a

CC pharmaceutical composition for inhibiting angiogenesis or for treating

CC cancer in a mammal. (I) Can be also used for the treatment or prevention

CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune

CC and degenerative arthritis. This is the amino acid sequence of an

CC angiogenesis inhibiting peptide of the invention.

XX

SQ Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;

Best Local Similarity 100.0%; Pred. NC. 1.4e+06;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GVITXIRP 9

Db 1 GVITXIRP 8

RESULT 11

ADD31269

ID ADD31269 standard; peptide; 8 AA.

XX

XX ADD31269;

XX

XX 15-JAN-2004 (first entry)

XX

XX Angiogenesis inhibiting peptide #44.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;

KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;

XX

XX

XX

XX

XX

XX

XX

XX



KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1 /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT Modified-site 5 /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT Misc-difference 6 /note= "D-form residue"  
 FT Modified-site 8 /note= "NHCH2CH3"

XX US2003109455-A1.

XX 12-JUN-2003.

XX 30-OCT-2002; 2002US-00283550.

XX 31-OCT-2001; 2001US-0335017P.

XX (HAI/) HAVIV F.  
 XX (BRAD/) BRADLEY M F.

XX Haviv F, Bradley MF;

XX WPI; 2003-843101/78.

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.

XX Claim 11; Page 24; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 1 GVITXIRP 8

RESULT 12

ADD31267  
 ID ADD31267 standard; peptide; 8 AA.

XX ADD31267;

XX 15-JAN-2004 (first entry)

XX Angiogenesis inhibiting peptide #42.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1 /label= OTHER  
 FT /note= "OTHER= N-(6-Me-nicotinyl)"  
 FT Misc-difference 3 /note= "D-form residue"  
 FT Modified-site 5 /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT Modified-site 8 /note= "NHCH2CH3"

XX US2003109455-A1.

XX 12-JUN-2003.

XX 30-OCT-2002; 2002US-00283550.

XX 31-OCT-2001; 2001US-0335017P.

XX (HAI/) HAVIV F.  
 XX (BRAD/) BRADLEY M F.

XX Haviv F, Bradley MF;

XX WPI; 2003-843101/78.

XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.

XX Claim 11; Page 24; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 1 GVITXIRP 8

RESULT 13

ADD31282  
 ID ADD31282 standard; peptide; 8 AA.

XX ADD31282;

XX 15-JAN-2004 (first entry)

XX Angiogenesis inhibiting peptide #57.

XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.

XX OS	Synthetic.	
XX FH	Key	Location/Qualifiers
XX FT	Modified-site	1 /label= OTHER
XX FT	Misc-difference	3 /note= "OTHER= N-acetyl"
XX FT	Misc-difference	3 /note= "D-form residue"
XX FT	Modified-site	8 /note= "NHCH2CH3"
XX FT	Modified-site	8 /note= "NHCH2CH3"
XX PN	US2003109455-A1.	
XX PD	12-JUN-2003.	
XX PF	30-OCT-2002; 2002US-00283550.	
XX PR	31-OCT-2001; 2001US-0335017P.	
XX PA	(HAVI/) HAVIV F.	
XX PA	(BRAD/) BRADLEY M F.	
XX PI	Haviv F, Bradley MF;	
XX PX	WPI; 2003-843101/78.	
XX PT	New hepta-, octa-, or nonapeptide compounds useful for inhibiting	
XX PT	angiogenesis, treating cancer in mammal, or for prevention of other	
XX PT	diseases such as autoimmune diseases.	
XX PS	Claim 13; Page 25; 26pp; English.	
XX PX	The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)	
XX CC	has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and	
XX CC	antiarthritic activities. (I) or a salt of it, is used in a	
XX CC	pharmaceutical composition for inhibiting angiogenesis or for treating	
XX CC	cancer in a mammal. (I) Can be also used for the treatment or prevention	
XX CC	of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune	
XX CC	and degenerative arthritis. This is the amino acid sequence of an	
XX CC	angiogenesis inhibiting peptide of the invention.	
XX SQ	Sequence 8 AA;	
XX	Query Match	91.9%; Score 34; DB 7; Length 8;
XX	Best Local Similarity	87.5%; Pred. No. 1.4e+06;
XX	Matches	7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	2 GVITXIRP 9	
DB	1 GVITQIRP 8	
XX	RESULT 14	
XX	ADE83640	
XX	ID ADE83640 standard; peptide; 8 AA.	
XX AC	ADE83640;	
XX DT	29-JAN-2004 (first entry)	
XX DE	Antiangiogenic peptide.	
XX KW	antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;	
XX KW	antiarthritic; ophthalmological; antidiabetic; dermatological;	
XX KW	antipsoriatic; antiarteriosclerotic; cardiant; vulnery;	
XX KW	antiinflammatory; antitumor; haemostatic; cancer; solid tumour;	
XX KW	autoimmune disease; rheumatoid arthritis; immune arthritis;	
XX KW	degenerative arthritis; ocular disease; skin disease;	
XX KW	blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;	
XX KW	plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;	
XX KW	wound granulation; birth control; angiogenesis; cat scratch disease;	
XX KW	ulcer.	
XX OS	Synthetic.	
XX FH	Key	Location/Qualifiers
XX FT	Modified-site	1 /note= "acetylated"
XX FT	Misc-difference	3 /note= "D-form residue"
XX FT	Modified-site	5 /label= Nva
XX FT	Misc-difference	6 /note= "norvaline"
XX FT	Misc-difference	8 /note= "D-form residue"
XX FT	Modified-site	8 /note= "Pro is C-terminally modified with -NHCH2CH3"
XX PN	WO2003037268-A2.	
XX PD	08-MAY-2003.	
XX PF	30-OCT-2002; 2002WO-US034811.	
XX PR	31-OCT-2001; 2001US-00000681.	
XX PR	04-OCT-2002; 2002US-00263812.	
XX PX	(ABBO ) ABBOTT LAB.	
XX PI	Haviv F, Bradley MF;	
XX PX	WPI; 2003-617886/58.	
XX PT	New hepta-, octa- and nona-peptide compounds used for treating e.g.	
XX PT	cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic	
XX PT	retinopathy.	
XX PX	Claim 11; Page 45; 51pp; English.	
XX CC	The present invention describes hepta-, octa- and nona-peptide compounds	
XX CC	(I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,	
XX CC	immunosuppressive, antirheumatic, antiarthritic, ophthalmological,	
XX CC	antidiabetic, dermatological, antipsoriatic, antiarteriosclerotic,	
XX CC	cardiant, vulnery, antiinflammatory, antitumor and haemostatic	
XX CC	activities. (I) can be used for treating cancer including primary and	
XX CC	metastatic solid tumours. (I) can also be used for treating autoimmune	
XX CC	diseases such as rheumatoid, immune and degenerative arthritis, ocular	
XX CC	diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,	
XX CC	myocardial angiogenesis, plaque neovascularisation, telangiectasia,	
XX CC	haemophilia, joints, angiofibroma and wound granulation, and for the	
XX CC	treatment of diseases involving excessive or abnormal stimulation of	
XX CC	endothelial cells. (I) can also be used as birth control agents and for	
XX CC	the treatment of diseases involving angiogenesis as a pathologic	
XX CC	consequence such as cat scratch disease and ulcers, and to reduce	
XX CC	bleeding by administration prior to surgery such as for the treatment of	
XX CC	tumours. The present sequence represents an antiangiogenic peptide, which	
XX CC	is used in the exemplification of the present invention.	
XX SQ	Sequence 8 AA;	
XX	Query Match	91.9%; Score 34; DB 7; Length 8;
XX	Best Local Similarity	100.0%; Pred. No. 1.4e+06;
XX	Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	2 GVITXIRP 9	
DB	1 GVITXIRP 8	
XX	RESULT 15	
XX	ADE83694	
XX	ID ADE83694 standard; peptide; 8 AA.	
XX AC	ADE83694;	
XX XX		

DT 29-JAN-2004 (first entry)  
 XX Antiangiogenic peptide.  
 DE  
 XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vulnary;  
 KW antiinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.  
 XX  
 OS Synthetic.

XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "acetylated"  
 FT Modified-site 2  
 FT /note= "N-methylvalyl"  
 FT Misc-difference 3  
 FT /note= "D-form residue"  
 FT Modified-site 5  
 FT /label= Nva  
 FT /note= "norvaline"  
 FT Modified-site 8  
 FT /note= "Pro is C-terminally modified with -NHCH2CH3"  
 XX  
 PN WO2003037268-A2.  
 XX  
 XX 08-MAY-2003.  
 XX  
 XX 30-OCT-2002; 2002WO-US034811.  
 XX  
 XX 31-OCT-2001; 2001US-00000681.  
 PR 04-OCT-2002; 2002US-00263912.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Bradley MF;  
 PI WPI; 2003-617886/58.  
 XX  
 DR New hepta-, octa- and nona-peptide compounds used for treating e.g.  
 XX cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic  
 XX retinopathy.  
 XX  
 PS Claim 11; Page 45; 51pp; English.

XX The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antiarteriosclerotic,  
 CC cardiant, vulnary, antiinflammatory, antiulcer and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumours. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.  
 XX  
 XX Sequence 8 AA;

Query Match 91.9%; Score 34; DB 7; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8, Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 GVITXIRP 9  
 Db 1 GVITXIRP 8  
 Search completed: April 1, 2004, 17:35:22  
 Job time : 44.1053 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:30:38 ; Search time 12.7895 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-2  
Perfect score: 37  
Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	86.5	1208	4	US-09-540-236-3604
2	31	83.8	748	4	Sequence 3604, Ap
3	29	78.4	352	4	Sequence 21696, A
4	28	75.7	140	4	Sequence 8196, Ap
5	28	75.7	240	2	Sequence 2611, Ap
6	28	75.7	240	2	Sequence 6, Appli
7	28	75.7	240	3	Sequence 8, Appli
8	28	75.7	240	3	Sequence 12, Appl
9	28	75.7	252	4	Sequence 14, Appl
10	28	75.7	323	4	Sequence 8187, Ap
11	28	75.7	441	4	Sequence 7409, Ap
12	28	75.7	551	4	Sequence 6369, Ap
13	28	75.7	727	4	Sequence 16, Appl
14	28	75.7	1048	4	Sequence 11, Appl
15	28	75.7	1131	4	Sequence 31247, A
16	28	75.7	1495	4	Sequence 1, Appli
17	28	75.7	1495	4	Sequence 1, Appli
18	27	73.0	19	4	Sequence 7, Appli
19	27	73.0	19	5	Sequence 4, Appli
20	27	73.0	23	1	PCT-US93-01652-4
21	27	73.0	23	2	Sequence 6, Appli
22	27	73.0	23	4	Sequence 6, Appli
23	27	73.0	23	4	Sequence 12, Appl
24	27	73.0	23	6	Sequence 12, Appl
25	27	73.0	40	2	Patent No. 5426100
26	27	73.0	59	4	Sequence 5, Appli
27	27	73.0	60	1	Sequence 18, Appl
28	27	73.0	60	1	Sequence 12, Appl
29	27	73.0	60	2	Sequence 12, Appl

28	27	73.0	60	6	5426100-12	Patent No. 5426100
29	27	73.0	76	4	US-09-252-991A-25888	Sequence 25888, A
30	27	73.0	83	4	US-09-621-976-4505	Sequence 4505, Ap
31	27	73.0	227	2	US-08-698-407-1	Sequence 1, Appli
32	27	73.0	227	3	US-09-195-855-1	Sequence 1, Appli
33	27	73.0	239	5	PCT-US93-01652-1	Sequence 1, Appli
34	27	73.0	308	4	US-09-352-991A-28762	Sequence 28762, A
35	27	73.0	318	4	US-09-134-000C-4662	Sequence 4662, Ap
36	27	73.0	321	4	US-09-198-452A-570	Sequence 570, App
37	27	73.0	381	2	US-08-858-052-3	Sequence 1, Appli
38	27	73.0	381	2	US-08-858-052-3	Sequence 3, Appli
39	27	73.0	381	3	US-09-200-284-1	Sequence 3, Appli
40	27	73.0	381	3	US-09-200-284-3	Sequence 3, Appli
41	27	73.0	413	4	US-08-134-001C-3702	Sequence 3702, Ap
42	27	73.0	426	4	US-09-328-352-6416	Sequence 6416, Ap
43	27	73.0	448	4	US-09-361-443-4	Sequence 4, Appli
44	27	73.0	467	4	US-09-361-443-2	Sequence 2, Appli
45	27	73.0	469	4	US-09-198-452A-793	Sequence 793, App

ALIGNMENTS

RESULT 1  
US-09-540-236-3604  
; Sequence 3604, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATP  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 3604  
; LENGTH: 1208  
; TYPE: PRT  
; ORGANISM: M.cattarrhalis  
US-09-540-236-3604

Query Match 86.5%; Score 32; DB 4; Length 1208;  
Best Local Similarity 62.5%; Pred. No. 51;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVTIXIRP 9  
Db 1113 GIITQVRP 1120

RESULT 2  
US-09-252-991A-21696  
; Sequence 21696, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21696  
; LENGTH: 748  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21696

Query Match 83.8%; Score 31; DB 4; Length 748;

```
Best Local Similarity 75.0%; Pred. No. 52;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 740 GVTHIRP 747

RESULT 3
US-09-328-352-8196
; Sequence 8196, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8196
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8196

Query Match 78.4%; Score 29; DB 4; Length 352;
Best Local Similarity 62.5%; Pred. No. 69;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 109 GIITAIHP 116

RESULT 4
US-09-540-236-2611
; Sequence 2611, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARR
; FILE REFERENCE: 2709 2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2611
; LENGTH: 140
; TYPE: PRT
; ORGANISM: M. catarrhalis
US-09-540-236-2611

Query Match 75.7%; Score 28; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 42 GVIMKIRP 49

RESULT 5
US-08-114-555A-6
; Sequence 6, Application US/08114555A
; Patent No. 5854392
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100
; NUMBER OF SEQUENCES: 18
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/114,555A
; FILING DATE: 30-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-114-555A-6

Query Match 75.7%; Score 28; DB 2; Length 240;
Best Local Similarity 62.5%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 3 GVITSLRP 10

RESULT 6
US-08-114-555A-8
; Sequence 8, Application US/08114555A
; Patent No. 5854392
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/114,555A
; FILING DATE: 30-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-115
```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-114-555A-8

Query Match 75.7%; Score 28; DB 2; Length 240;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|||:|  
Db 3 GVTTSLRP 10

RESULT 7  
US-08-559-397A-12  
Sequence 12, Application US/08559397A  
Patent No. 6083713  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.  
APPLICANT: Kozlowski, Michael R.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF  
TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,397A  
FILING DATE: 15-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-559-397A-12

Query Match 75.7%; Score 28; DB 3; Length 240;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|||:|  
Db 3 GVTTSLRP 10

RESULT 8  
US-08-559-397A-14  
Sequence 14, Application US/08559397A  
Patent No. 6083713  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.  
APPLICANT: Kozlowski, Michael R.  
APPLICANT: Neve, Rachael L.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF  
TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,397A  
FILING DATE: 15-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-559-397A-14

Query Match 75.7%; Score 28; DB 3; Length 240;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|||:|  
Db 3 GVTTSLRP 10

RESULT 9  
US-09-328-352-8187  
Sequence 8187, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQUENCE CHARACTERISTICS:  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-8187

```
Query Match      75.7%; Score 28; DB 4; Length 252;
Best Local Similarity 50.0%; Pred. No. 82;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      74 GVVAVRP 81

RESULT 10
US-09-489-039A-7409
; Sequence 7409, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7409
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7409

Query Match      75.7%; Score 28; DB 4; Length 323;
Best Local Similarity 50.0%; Pred. No. 1,1e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      119 GVVSALRP 126

RESULT 11
US-09-328-352-6369
; Sequence 6369, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 5252
; SEQ ID NO 6369
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6369

Query Match      75.7%; Score 28; DB 4; Length 441;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      361 GPITSVRP 368

RESULT 12
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
```

```
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-16

Query Match      75.7%; Score 28; DB 4; Length 551;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      313 GIVTSTRP 320

RESULT 13
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inosuchi, Ei-ji
; APPLICANT: Hakoziaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE OF INVENTION: composition and method of immunologically analyzing human ADAMT
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match      75.7%; Score 28; DB 4; Length 727;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      489 GIVTSTRP 496

RESULT 14
US-09-921-099A-11
; Sequence 11, Application US/09921099A
; Patent No. 6602707
; GENERAL INFORMATION:
; APPLICANT: Hefeneider, Steven
; APPLICANT: Merkins, Louise
; APPLICANT: Bennett, Robert
; APPLICANT: Seiss, Donald
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene
; FILE REFERENCE: 00-617-A
; CURRENT APPLICATION NUMBER: US/09/921,099A
; CURRENT FILING DATE: 2001-08-01
```

```

; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent version 3.0
; SEQ ID NO 11
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-921-099A-11

Query Match      75.7%; Score 28; DB 4; Length 1048;
Best Local Similarity 62.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      957 GULTPIRP 964

RESULT 15
US-09-252-991A-31247
; Sequence 31247, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31247
; LENGTH: 1131
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31247

Query Match      75.7%; Score 28; DB 4; Length 1131;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVITXIRP 9
Db      1054 GAITRLRP 1061

Search completed: April 1, 2004, 17:42:11
Job time : 13.7895 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 9.78947 Seconds  
(without alignments)  
88.434 Million cell updates/sec

Title: US-09-833-196-2

Perfect score: 37  
Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: Pirl1:\*

2: Pirl2:\*

3: Pirl3:\*

4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	86.5	141	2 S09804	Hypothetical prote
2	32	86.5	496	2 G87546	acid-CoA ligase, p
3	31	83.8	166	2 A69820	hypothetical prote
4	31	83.8	288	2 A11956	permease protein o
5	31	83.8	401	2 A69713	septal wall dissol
6	31	83.8	498	2 H83733	hypothetical prote
7	30	81.1	77	2 AD0343	hypothetical prote
8	30	81.1	355	2 D97339	dioxygenase relate
9	29	78.4	212	2 H87472	hypothetical prote
10	29	78.4	377	2 C82876	oligopeptide trans
11	29	78.4	391	2 T36739	hypothetical prote
12	29	78.4	471	2 B33828	aspartate ammonia-
13	29	78.4	475	1 UPESD	aspartate ammonia-
14	29	78.4	2685	2 T38755	hypothetical prote
15	29	78.4	4307	2 T20721	hypothetical prote
16	28	75.7	152	2 B70467	conserved hypothet
17	28	75.7	287	2 C81066	conserved hypothet
18	28	75.7	300	2 A81804	probable integral
19	28	75.7	302	2 G86920	hypothetical prote
20	28	75.7	302	2 B70888	hypothetical prote
21	28	75.7	309	2 A81534	oxidoreductases ho
22	28	75.7	309	2 AF1176	oxidoreductases ho
23	28	75.7	318	2 AE0251	exported high-affi
24	28	75.7	327	1 S26693	transcription fact
25	28	75.7	340	2 F97251	shape-determining
26	28	75.7	369	2 F82696	rod shape-determin
27	28	75.7	437	2 C72570	probable translati
28	28	75.7	443	2 F83759	cytochrome d (bd-t
29	28	75.7	640	2 B84478	probable replicati

30 28 75.7 728 2 A70632  
31 28 75.7 919 2 P81998  
32 28 75.7 919 2 P81225  
33 28 75.7 951 2 T00017  
34 28 75.7 1048 2 T23764  
35 28 75.7 1051 2 A38373  
36 28 75.7 1051 2 B83330  
37 28 75.7 1067 2 D75625  
38 28 75.7 1375 1 J05148  
39 28 75.7 1495 2 S60255  
40 28 75.7 1551 2 AH3488  
41 28 75.7 2796 2 JC4743  
42 28 75.7 5032 1 A35041  
43 28 75.7 5035 1 I46646  
44 28 75.7 5037 2 B35041  
45 27 73.0 54 2 S09544

#### ALIGNMENTS

##### RESULT 1

S09804 Hypothetical protein UL41 - human cytomegalovirus (strain AD169)

C:Species: human cytomegalovirus, human herpesvirus 5

A:Note: host Homo sapiens (man)

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 28-Jul-2000

C:Accession: S09804

R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;

M.; Barrell, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirt

A:Reference number: S09749; MUID:90269039; PMID:2161319

A:Accession: S09804

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-141 <CHE>

A:Cross-references: EMBL:X17403; NID:G59591; PIDN:CAA35400.1; PID:e27240; PID:g1813967

A:Note: this sequence was submitted to the EMBL Data Library, December 1989

A:Note: this reading frame extends between two stop codons and does not begin with a st

C:Superfamily: human cytomegalovirus hypothetical protein UL41

Query Match 86.5%; Score 32; DB 2; Length 141;

Best Local Similarity 62.5%; Pred. No. 3.7;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVTXIRP 9

Db 8 GIITLRP 15

##### RESULT 2

G87546

acid-CoA ligase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C:Accession: G87546

R:Nierman, W.C.; Feldblym, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87546

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-496 <STO>

A:Cross-references: GB:AE005673; NID:g13423937; PIDN:AAK24371.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2400

C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match

86.5%; Score 32; DB 2; Length 496;

Best Local Similarity 75.0%; Pred. No. 14;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|:|  
Db 196 GLITSIRP 203

RESULT 3  
A69820  
hypothetical protein yhaw - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: A69820  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, K.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69820  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-166 <KUN>  
A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12823.1; PID:G2633319  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yhaw  
C:Superfamily: Bacillus subtilis hypothetical protein yhaw

Query Match 83.8%; Score 31; DB 2; Length 166;  
Best Local Similarity 62.5%; Pred. No. 7.7;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|:|  
Db 108 GLITGIRP 115

RESULT 4  
A11966  
permease protein of ABC transporter all1284 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A:Notes: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C:Accession: A11966  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: A11966  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-288 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BAB73241.1; PID:G17130631; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all1284  
C:Superfamily: leucine transport protein livH

Query Match 83.8%; Score 31; DB 2; Length 288;  
Best Local Similarity 62.5%; Pred. No. 14;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|:|  
Db 209 GLITAVRP 216

RESULT 5  
A69713  
septal wall dissolution protein spoIIP - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: A69713  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Togononi, K.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69713  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-401 <KUN>  
A:Cross-references: GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14495.1; PID:G2634996  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: spoIIP

Query Match 83.8%; Score 31; DB 2; Length 401;  
Best Local Similarity 62.5%; Pred. No. 19;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|:|  
Db 38 GVLTSIRP 45

RESULT 6  
H83793  
hypothetical protein BH152 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: H83793  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans an  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: H83793  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-498 <STO>  
A:Cross-references: GB:AP001511; GB:BA000004; NID:G10173727; PIDN:BAB04871.1; GSPDB:GN  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH152

Query Match 83.8%; Score 31; DB 2; Length 498;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|:|  
Db 107 GVLTVGRP 114

RESULT 7  
AD0343  
hypothetical protein YP02818 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AD0343  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AD0343  
A:Molecule type: DNA  
A>Status: preliminary  
A:Residues: 1-77 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC93051.1; PID:gt15980789; GSPDB:GN00175  
C:Genetics:  
A:Gene: YP02818

Query Match 81.1%; Score 30; DB 2; Length 77;  
Best Local Similarity 62.5%; Pred. No. 6;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9  
|||||  
Db 59 GVITSVQP 66

RESULT 8  
D97339  
dioxigenase related to 2-nitropropane dioxigenase [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: D97339  
R:Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: D97339  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-355 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK81503.1; PID:gt5026676; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3580  
C:Superfamily: Campylobacter jejuni hypothetical protein Cj1270c

Query Match 81.1%; Score 30; DB 2; Length 355;  
Best Local Similarity 62.5%; Pred. No. 30;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9  
|||||  
Db 190 GVITAKP 197

RESULT 9  
H87472  
hypothetical protein CC1804 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87472  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87429; MUID:21173698; PMID:11259647

A:Accession: H87472  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-212 <SAO>  
A:Cross-references: GB:AB005673; NID:gl3423238; PIDN:AAK23780.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC1804

Query Match 78.4%; Score 29; DB 2; Length 212;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9  
|||||  
Db 30 GVITLARP 37

RESULT 10  
C82876  
oligopeptide transport system permease protein UUS62 [imported] - Ureaplasma urealyticu  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82876  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
A:Accession: C82876  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-377 <GLA>  
A:Cross-references: GB:AB002154; GB:AF222894; NID:96899557; PIDN:AAF30975.1; GSPDB:GN00  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: oppB; UUS62  
A:Genetic code: SGC3

Query Match 78.4%; Score 29; DB 2; Length 377;  
Best Local Similarity 75.0%; Pred. No. 55;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9  
|||||  
Db 146 GVSTIRP 153

RESULT 11  
T36739  
hypothetical protein SC111.06 - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T36739  
R:Saunders, D.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, July 1999  
A:Reference number: Z21613  
A:Accession: T36739  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-391 <SAU>  
A:Cross-references: EMBL:AL096849; PIDN:CAB50932.1; GSPDB:GN00070; SCOEDB:SC111.06  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC111.06

Query Match 78.4%; Score 29; DB 2; Length 391;  
Best Local Similarity 62.5%; Pred. No. 57;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9  
|||||  
Db 309 GTLTAIRP 316

RESULT 12  
B83828  
C:Species: aspartate ammonia-lyase ansB [imported] - Bacillus halodurans (strain C-125)  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: B83828  
R: Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A89580; MUID:20512582; PMID:11058132  
A:Accession: B83828  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-471 <GTO>  
A:Cross-references: GB:AP001512; GB:BA000004; NID:G10174030; PIDN:BA805145.1; GSPDB:GN000004  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: ansB  
C:Superfamily: fumarate hydratase

Query Match 78.4%; Score 29; DB 2; Length 471;  
Best Local Similarity 62.5%; Pred. No. 69;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|  
Db 406 GVITAINP 413

RESULT 13  
UFBS0  
aspartate ammonia-lyase (EC 4.3.1.1) - Bacillus subtilis  
N:Alternate names: L-asparaginase (ansB)  
C:Species: Bacillus subtilis  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 16-Jun-2000  
C:Accession: B39440; A69586  
R: Sun, D.; Setlow, P.  
J. Bacteriol. 173, 3831-3845, 1991  
A:Title: Cloning, nucleotide sequence, and expression of the Bacillus subtilis ans operon  
A:Reference number: A39440; MUID:91267950; PMID:1711029  
A:Accession: B39440  
A:Molecule type: DNA  
A:Residues: 1-475 <SUN>  
A:Cross-references: GB:M63264; NID:G142516; PIDN:AAA22244.1; PID:G142518  
R: Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tostato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69586  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-475 <KUN>  
A:Cross-references: GB:Z99116; GB:AL009126; NID:G2634723; PIDN:CAB14289.1; PID:G2634792  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: ansB  
C:Superfamily: fumarate hydratase  
C:Keywords: ammonia-lyase; carbon-nitrogen lyase

Query Match 78.4%; Score 29; DB 1; Length 475;  
Best Local Similarity 62.5%; Pred. No. 70;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|:|:|:|  
Db 410 GVITAVNP 417

RESULT 14  
T38755  
hypothetical protein SPAC3H5.09c - fission yeast (schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T38755  
R: Oliver, K.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z21809  
A:Accession: T38755  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2685 <OLJ>  
A:Cross-references: EMBL:Z99296; PIDN:CAB16593.1; GSPDB:GN00066; SPDB:SPAC3H5.09c  
A:Experimental source: strain 972h-; cosmid C3H5  
C:Genetics:  
A:Gene: SPDB:SPAC3H5.09c  
A:Map position: 1

Query Match 78.4%; Score 29; DB 2; Length 2685;  
Best Local Similarity 85.7%; Pred. No. 4.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 VITXIRP 9  
|:|:|:|  
Db 625 VITAIRP 631

RESULT 15  
T20721  
hypothetical protein F25F2.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 19-May-2000  
C:Accession: T20721; T21343; T23842  
R: Ainscough, R.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19314  
A:Accession: T20721  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4307 <WIL>  
A:Cross-references: EMBL:Z35662; PIDN:CAA84721.1; GSPDB:GN00021; CESP:F25F2.2  
R: Ainscough, R.  
submitted to the EMBL Data Library, August 1994  
A:Reference number: Z19410  
A:Accession: T21343  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4307 <W12>  
A:Cross-references: EMBL:Z35599; PIDN:CAA84661.1; GSPDB:GN00021; CESP:F25F2.2  
R: Sulston, J.  
submitted to the EMBL Data Library, June 1994  
A:Reference number: Z19806  
A:Accession: T23842  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-4307 <W13>  
A:Cross-references: EMBL:Z34802; PIDN:CAA84339.1; GSPDB:GN00021; CESP:F25F2.2  
A:Experimental source: clone M88  
C:Genetics:  
A:Gene: CESP:F25F2.2  
A:Map position: 3  
A:Introns: 79/3; 167/3; 251/3; 328/3; 587/3; 625/3; 737/3; 1006/1; 1190/2; 1494/2; 1770  
C:Superfamily: rat MEGF1 protein; cadherin repeat homology; EGF homology; laminin G ref

F:3761-3900/Domain: laminin G repeat homology <LGR>

Query Match 78.4%; Score 29; DB 2; Length 4307;  
Best Local Similarity 50.0%; Pred. No. 7e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
||| :|  
Db 3045 GWTIVVKP 3052

Search completed: April 1, 2004, 17:40:38  
Job time : 11.7895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:18 ; Search time 5.52632 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-2

Perfect score: 37

Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	32	86.5	141	1 UL41_HCMVA	P16814 human cytom
2	31	83.8	401	1 SP2P_BACSU	P27968 bacillus su
3	29	78.4	375	1 ARI1_HUMAN	Q96a54 homo sapien
4	29	78.4	375	1 ARI1_MOUSE	Q91v11 mus musculu
5	29	78.4	475	1 ASPA_BACSU	P26899 bacillus su
6	28	75.7	126	1 YJ53_AQUAE	O57766 aquifex ae
7	28	75.7	198	1 IPE2_ARATH	O44663 arabidopsis
8	28	75.7	318	1 ZNUA_YERPE	Q8zeu2 yersinia pe
9	28	75.7	352	1 YORM_TTV1	P19297 thermoprote
10	28	75.7	370	1 PIT1_MELGA	Q05749 melagris g
11	28	75.7	379	1 OXAA_PROMP	Q7v0r8 prochloroco
12	28	75.7	437	1 EFA1_AERPE	Q9yav0 aeropyrum p
13	28	75.7	512	1 Y4UC_RHISN	Q53197 rhizobium s
14	28	75.7	968	1 AT51_MOUSE	P37857 mus musculu
15	28	75.7	1051	1 UBA1_WHEAT	P20973 triticum ae
16	28	75.7	1051	1 UBA2_WHEAT	P31251 triticum ae
17	28	75.7	2472	1 NCR2_MOUSE	Q9w442 mus musculu
18	28	75.7	2517	1 NCR2_HUMAN	Q9y618 h nuclear r
19	28	75.7	5035	1 RYR1_PIG	P16960 sus scrofa
20	28	75.7	5037	1 RYR1_RABIT	P11716 cryptolagus
21	28	75.7	5038	1 RYR1_HUMAN	P21817 homo sapien
22	27	73.0	54	1 VRP1_BPT3	P20836 bacterioph
23	27	73.0	154	1 CRB1_STRCO	Q9fc39 streptomyce
24	27	73.0	177	1 TRF4_ECOLI	Q93450 escherichia
25	27	73.0	277	1 PSB0_SYNE7	P11472 synchococc
26	27	73.0	295	1 PNK1_PSEAE	Q9hz20 pseudomonas
27	27	73.0	296	1 PNK1_PSEPK	Q881c3 pseudomonas
28	27	73.0	296	1 PNK1_PSESM	Q87y42 pseudomonas
29	27	73.0	301	1 YA93_MYCPN	P75559 mycoplasma
30	27	73.0	401	1 YQV1_CABEL	Q95551 caenorhabdi
31	27	73.0	404	1 METK_SULSO	Q98099 sulfolobus
32	27	73.0	405	1 METK_SULTO	Q976f3 sulfolobus
33	27	73.0	406	1 METK_AERPE	Q9ybk2 aeropyrum p

34	27	73.0	423	1 GLGC_THEMEA	Q9wv82 thermotoga
35	27	73.0	436	1 DNAA_THETH	Q9x9d5 thermus the
36	27	73.0	467	1 NORA_CHLPN	Q927g2 chlamydia p
37	27	73.0	500	1 LIPE_MOUSE	Q9wv95 mus musculu
38	27	73.0	515	1 DNAA_MICLU	P21173 micrococci
39	27	73.0	524	1 DNAA_CORGL	Q9nu48 corynebacte
40	27	73.0	545	1 TX32_CABEL	O61764 caenorhabdi
41	27	73.0	573	1 DNAA_COREF	O8ful7 corynebacte
42	27	73.0	759	1 GSPD_XANCP	P29041 xanthomonas
43	27	73.0	882	1 CAD1_HUMAN	P12830 homo sapien
44	27	73.0	902	1 EPBB_XENLA	Q91736 xenopus lae
45	27	73.0	974	1 EPB3_XENLA	Q91735 xenopus lae

#### ALIGNMENTS

##### RESULT 1

UL41_HCMVA	UL41_HCMVA	STANDARD;	PRT;	141 AA.
ID	P16814;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	01-FEB-1991 (Rel. 17, Last annotation update)			
DE	Hypothetical protein UL41.			
GN	UL41.			
OS	Human cytomegalovirus (strain AD169).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
OX	NCBI TaxID=10360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90269039; PubMed=2161319;			
RA	Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,			
RA	Horsnell T., Hutchinson C.A. III, Kourzides T., Martignetti J.A.,			
RA	Predelle E., Satchwell S.C., Tomlinson P., Weston K.W., Barrell B.G.;			
RT	"Analysis of the protein-coding content of the sequence of human			
RT	Cytomegalovirus strain AD169."			
RL	Curr. Top. Microbiol. Immunol. 154:125-169(1990).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL; X17403; GAA35400.1; -			
DR	PIR; S09804; S09804.			
KW	Hypothetical protein.			
SC	SEQUENCE 141 AA; 16766 MW; 20005377B1EFB712 CRC64;			

Query Match 86.5%; Score 32; DB 1; Length 141;  
Best Local Similarity 62.5%; Pred. No. 3;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY	2	GVITXIRP 9
DB	8	GIITLIRP 15

##### RESULT 2

SP2P_BACSU	SP2P_BACSU	STANDARD;	PRT;	401 AA.
ID	P37568;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Stage II sporulation protein P.			
GN	SPOIIP OR BSU25530.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			

```

OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Takemaru K.I., Sato T., Kobayashi Y.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DDBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-332 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Knevedo V., Berto M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.K., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigic C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujita C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Rieger M., Rivolta C., Ruelle B., Rapoport G., Rey M., Reynolds S.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpsira P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wibatt A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
[4]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@sib-sib.ch).
[5]
CC EMBL: D17650; BAA04542.1; -
CC EMBL: D84432; BAA12458.1; -
CC EMBL: Z99117; CAB14495.1; -
CC F01; A69713; A69713.
CC Subtilisin; SG10439; spoiP.
KW Sporulation; Complete proteome.
SQ SEQUENCE 401 AA; 44548 MW; EB060014088E17A5 CRC64;
Query Match 93.8%; Score 31; DB 1; Length 401;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
2 GVITXRP 9

```

```

Db 38 GVITSLRP 45
||:|:|:|
RESULT 3
ADRI1 HUMAN
ID ADRI1 HUMAN STANDARD; PRT; 375 AA.
AC Q96A54; Q9Y360;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Adiponectin receptor protein 1 (CGI-45).
GN ADIPOR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RT Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
[2]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loughran J.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltan E., Kettner M., Maman A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[3]
RP FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=22687101; PubMed=12802337;
RA Sugiyama T., Miyagishi M., Hara K., Tsunoda M., Murakami K.,
RA Ohteki T., Uchida S., Takekawa S., Waki H., Tsuno N.H., Shibata Y.,
RA Terauchi Y., Froquel P., Tobe K., Koyasu S., Taira K., Kitamura T.,
RA Shimizu T., Nagai R., Kadowaki T.;
RT "Cloning of adiponectin receptors that mediate antidiabetic metabolic
RT effects.";
RL Nature 413:762-769(2003).
[4]
CC -!- FUNCTION: Receptor for globular and full-length adiponectin
CC (APM1), an essential hormone secreted by adipocytes that acts as
CC an antidiabetic. Probably involved in metabolic pathways that
CC regulate lipid metabolism such as fatty acid oxidation. Mediates
CC increased AMPK, PPARA ligand activity, fatty acid oxidation and
CC glucose uptake by adiponectin. Has some high-affinity receptor for
CC globular adiponectin but low-affinity receptor for full-length
CC adiponectin.
CC -!- SUBUNIT: May form homo and heteromultimers.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized to the
CC cell membrane and intracellular organelles.
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in skeletal
CC muscle. Expressed at intermediate level in brain, heart, spleen,
CC kidney, liver, placenta, lung and peripheral blood leukocytes.
CC weakly expressed in colon, thymus and small intestine.

```

```

CC -!- SIMILARITY: Belongs to the ADIPOR family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 369.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF151803; AAD34040.1; ALT_FRAME.
CC EMBL; BC001594; AAH01594.1; -.
CC EMBL; BC010743; AAH10743.1; -.
CC MIM; 607945; -.
CC DR InterPro; IPR004254; HlyIII-related.
CC DR Pfam; PF03006; UPF0073; 1.
CC KW Fatty acid metabolism; Lipid metabolism; Receptor; Transmembrane.
CC FT DOMAIN 1 136 CYTOPLASMIC (POTENTIAL).
CC FT TRANSNEM 137 157 1 (POTENTIAL).
CC FT DOMAIN 158 170 2 (POTENTIAL).
CC FT TRANSNEM 171 191 2 (POTENTIAL).
CC FT DOMAIN 192 209 3 (POTENTIAL).
CC FT TRANSNEM 210 230 3 (POTENTIAL).
CC FT DOMAIN 231 235 4 (POTENTIAL).
CC FT TRANSNEM 236 256 4 (POTENTIAL).
CC FT DOMAIN 257 267 5 (POTENTIAL).
CC FT TRANSNEM 268 288 5 (POTENTIAL).
CC FT DOMAIN 289 296 6 (POTENTIAL).
CC FT TRANSNEM 297 317 6 (POTENTIAL).
CC FT DOMAIN 318 331 7 (POTENTIAL).
CC FT TRANSNEM 332 352 7 (POTENTIAL).
CC FT DOMAIN 353 375 EXTRACELLULAR (POTENTIAL).
CC SEQUENCE 375 AA; 42615 MW; 1CC0300A7D178EB0 CRC64;
CC -----
Query Match 78.48; Score 29; DB 1; Length 375;
Best Local Similarity 50.08; Pred. No. 37;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
CC -----
Cc 2 GVITXIRP 9
Db 152 GILTLWRP 159
CC -----
RESULT 4
ADRI MOUSE
AC Q91VHL: 09CZAO; STANDARD; PRT; 375 AA.
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
GN ADIPOR1.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_Taxid=10090;
CC -----
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Satoh K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaï I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

```

```

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seyer T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Yawshav-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -----
RN SEQUENCE FROM N.A.
RP TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Trnchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -----
RP FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=22687101; PubMed=12802337;
RA Sugiyama T., Miyagishi M., Hara K., Tsunoda M., Murakami K.,
RA Ohteki T., Uchida S., Takekawa S., Waki H., Tsuno N.H., Shibata Y.,
RA Terauchi Y., Froquel P., Tobe K., Koyasu S., Taira K., Kitamura T.,
RA Shimizu T., Nagai R., Kadowaki T.;
RT "Cloning of adiponectin receptors that mediate antidiabetic metabolic
RL effects.";
RL Nature 423:762-769(2003).
CC -!- FUNCTION: Receptor for globular and full-length adiponectin
CC (APM1), an essential hormone secreted by adipocytes that acts as
CC an antidiabetic. Probably involved in metabolic pathways that
CC regulate lipid metabolism such as fatty acid oxidation. Mediates
CC increased AMPK, PPARA ligand activity, fatty acid oxidation and
CC glucose uptake by adiponectin. Has some high-affinity receptor for
CC globular adiponectin but low-affinity receptor for full-length
CC adiponectin.
CC -!- SUBUNIT: May form homo and heteromultimers (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized to the
CC cell membrane and intracellular organelles (By similarity).
CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in heart, kidney,
CC liver, lung, skeletal muscle and spleen. Weakly expressed in brain
CC and testis.
CC -!- SIMILARITY: Belongs to the ADIPOR family.
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a stop
CC codon in position 353 which was translated as Tyr to extend the
CC sequence and to a frameshift in position 375, which abolishes the
CC stop codon.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AK012847; BAB28509.1; ALT_SEQ.
CC EMBL; BC014875; AAH14875.1; -.

```



```

DR MGD: MGI:1919924; 2810031111Rik.
DR InterPro: IPR004254; H4YIII-related.
DR Pfam: PF03006; UPR0073; 1.
KW Fatty acid metabolism; Lipid metabolism; Receptor; Transmembrane.
FT DOMAIN 1 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 158 170 2 (POTENTIAL).
FT TRANSMEM 171 191 2 (POTENTIAL).
FT DOMAIN 192 209 3 (POTENTIAL).
FT TRANSMEM 210 230 3 (POTENTIAL).
FT DOMAIN 231 235 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 236 256 4 (POTENTIAL).
FT DOMAIN 257 267 5 (POTENTIAL).
FT TRANSMEM 268 288 5 (POTENTIAL).
FT DOMAIN 289 296 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 297 317 6 (POTENTIAL).
FT DOMAIN 318 331 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 332 352 7 (POTENTIAL).
FT DOMAIN 353 375 EXTRACELLULAR (POTENTIAL).
SQ SEQUENCE 375 AA; 42366 MW; 0E72F81B5E9938CE CRC64;

Query Match 78.4%; Score 29; DB 1; Length 375;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
Db 152 GILTMLRP 159

RESULT 5
ID ASPA_BACSU STANDARD; PRT; 475 AA.
AC P26899;
DT 01-AUG-1992 (Rel. 23; Created)
DT 01-AUG-1992 (Rel. 23; Last sequence update)
DT 15-MAR-2004 (Rel. 43; Last annotation update)
DE Aspartate ammonia-lyase (EC 4.3.1.1) (Aspartase).
GN ANSB OR BSU23570.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91267950; PubMed=1711029;
RA Sun D., Setlow P.;
RT "Cloning, nucleotide sequence, and expression of the Bacillus
RT subtilis ans operon, which codes for L-asparaginase and
RT L-aspartase."
RL J. Bacteriol. 173:3831-3845(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=9124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.;
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes".
RL Microbiology 142:3103-3111 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boursier R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

```

```

RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kashiara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koster P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porcollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccioni B., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Toketo V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -1- CATALYTIC ACTIVITY: L-aspartate = fumarate + NH(3).
CC -1- SUBUNIT: Homotetramer.
CC -1- SIMILARITY: Belongs to the class-II fumarase / aspartase family.
CC Aspartase subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M63264; AAA22244.1; -
CC DR EMBL: D84432; BAA12643.1; -
CC DR EMBL: Z99116; CAB14289.1; -
CC DR PIR: B39440; UFBSD.
CC DR HSP: P04422; LJSW.
CC DR InterPro: BG10301; ansb.
CC DR InterPro: IPR004708; Aspa.
CC DR InterPro: IPR00362; Fumarate lyase.
CC DR InterPro: IPR008948; L-Aspartase-like.
CC DR Pfam: PF00206; Lyase 1; 1.
CC DR PRINTS: PR00149; FUMATELYASE.
CC DR TIGRFAM: TIGR00839; aspa; 1.
CC DR PROSITE: PS00163; FUMARATE_LYASES; 1.
CC KW Lyase; Complete proteome.
CC SEQUENCE 475 AA; 52553 MW; C75133B2AE5B996B CRC64;

Query Match 78.4%; Score 29; DB 1; Length 475;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
Db 410 GVITAVNP 417

RESULT 6
YJ53_AQUAE STANDARD; PRT; 126 AA.
ID YJ53_AQUAE
AC O67766;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Hypothetical protein AQ_1953.
GN AQ_1953.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
CC NCBI_TaxID=63363;
CX [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=9819666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., R.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.U., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC
CC -1- SIMILARITY: Belongs to the rusa family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000765; AAC07739.1; ALT_INIT.
DR InterPro; IPR008822; Rusa.
DR Pfam; PF05866; Rusa; 1.
KW Hypothetical protein; Hydrolase; Nuclease; Endonuclease;
KW Complete proteome.
SQ SEQUENCE 126 AA; 14778 MW; 05BC64FAA7B9F9B1 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 126;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
DB 115 GVITXIRP 122

RESULT 7
IFE2 ARATH
ID IPE2 ARATH STANDARD; PRT; 198 AA.
AC O04653; OSAN8;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Eukaryotic translation initiation factor 4E-2 (EIF4E-2) (EIF-4E-2)
DE (mRNA cap-binding protein) (EIF-(iso)4F 25 kDa subunit) (EIF-(iso)4F
DE p28 subunit) (EIF4E iso protein) (EIF(iso)4E).
DE EIF4E2 OR AT5G35620 OR MJ24.8.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCB1_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.
RX MEDLINE=97378268; PubMed=9234949;
RA Wittmann S., Chatel H., Fortin M.G., Laliberte J.F.;
RT "Interaction of the viral protein genome linked of turnip mosaic
RT potyvirus with the translational eukaryotic initiation factor (iso)
RT 4E of Arabidopsis thaliana using the yeast two-hybrid system."
RL Virology 234:84-92(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Rodriguez C.M., Freire M.A., Robaglia C.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RC MEDLINE=98403884; PubMed=9734815;
RA Korani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned P1 and TAC clones."
RL DNA Res. 5:203-216(1998).

```

```

CC -1- FUNCTION: Recognizes and binds the 7-methylguanosine-containing
CC mRNA cap during an early step in the initiation of protein
CC synthesis and facilitates ribosome binding by inducing the
CC unwinding of the mRNAs secondary structures.
CC -1- SUBUNIT: EIF4F is a multi-subunit complex, the composition of
CC which varies with external and internal environmental conditions.
CC It is composed of at least EIF4A, EIF4E and EIF4G. EIF4E is also
CC known to interact with other partners. In higher plants two
CC isoforms of EIF4F have been identified, named EIF4F and
CC EIF(iso)4F. EIF4F has subunits p220 and p28, whereas EIF(iso)4F
CC has subunits p82 and p26. This isoform interacts with the viral
CC protein genome linked (VPG)-proteinase of turnip mosaic potyvirus.
CC -1- TISSUE SPECIFICITY: Abundant in floral organs and in young
CC developing tissues.
CC -1- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U62044; AAB66906.1; ALT_INIT.
DR EMBL; Y10547; CAA71579.1; -.
DR HSP; P07260; 1AP8
DR InterPro; IPR001040; TIF_EIF_4E.
DR Pfam; PF01652; IF4E; 1.
DR ProDom; PD003697; TIF_EIF_4E; 1.
DR PROSITE; PS00813; IF4E; 1.
KW Protein biosynthesis; Translation regulation; Initiation factor;
KW RNA-binding; Multigene family.
SQ SEQUENCE 198 AA; 22514 MW; 71FEB309E073A9D2 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 198;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
DB 139 GVITXIRP 146

RESULT 8
ZNUA YERPE
ID ZNUA YERPE STANDARD; PRT; 318 AA.
AC O8ZET2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE High-affinity zinc uptake system protein znuA precursor.
DE ZNUA OR YPO2061 OR Y2249.
GN Yersinia pestis.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OC NCB1_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague."
RT Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.

```

STRAIN=KIM5 / Bivovar Mediaevalis;  
MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.B., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM.":  
RL J. Bacteriol. 184:4601-4611 (2002).  
CC -!- FUNCTION: Involved in the high-affinity zinc uptake transport  
CC system (By similarity).  
CC -!- SUBCELLULAR LOCATION: Periplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the bacterial solute-binding protein family  
CC 9.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AJ414151; CAC90873.1; -;  
CC EMBL; AE013827; AAM85809.1; -;  
CC PIR; AE0251; AE0251.1; -;  
CC InterPro; IPR006127; SPB\_bac\_9.  
CC Pfam; PF01297; SPB\_bac\_9; 1.  
CC Transport; Periplasmic; Zinc transport; Zinc; Metal-binding; signal;  
CC Complete proteome.  
CC SIGNAL 1 28 POTENTIAL.  
CC CHAIN 29 318 HIGH-AFFINITY ZINC UPTAKE SYSTEM PROTEIN  
CC ZNUA.  
CC FT DOMAIN 132 147 HIS-RICH.  
CC FT METAL 62 62 ZINC (BY SIMILARITY).  
CC FT METAL 153 153 ZINC (BY SIMILARITY).  
CC FT METAL 217 217 ZINC (BY SIMILARITY).  
CC SQ SEQUENCE 318 AA; 35244 MW; 441526855471F08C CRC64;  
-----  
CC Query Match 75.7%; Score 28; DB 1; Length 318;  
CC Best Local Similarity 71.4%; Pred. No. 54;  
CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
-----  
CC YQYM TTV1 9  
CC 3 VITXIRP 9  
CC 30 VTSIRP 36  
-----  
CC RESULT 9  
CC YQYM TTV1 STANDARD; PRT; 352 AA.  
CC AC P19297;  
CC DT 01-NOV-1990 (Rel. 16, Created)  
CC DT 01-NOV-1990 (Rel. 16, Last sequence update)  
CC DT 01-NOV-1990 (Rel. 16, Last annotation update)  
CC DE Hypothetical 38.6 kDa protein.  
CC OS Thermoproteus tenax virus 1 (strain KRA1) (TTV1).  
CC OC Viruses; dsDNA viruses, no RNA stage; Lipothrixviridae;  
CC OC Lipothrixvirus.  
CC OX NCBI\_TaxID=10480;  
CC [1]\_TaxID=10480;  
CC RN SEQUENCE FROM N.A.  
CC RA Neumann H.;  
CC RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; X14855; CAA32993.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 352 AA; 38674 MW; 0D0E49B9E08E6A88 CRC64;  
-----  
CC Query Match 75.7%; Score 28; DB 1; Length 352;  
CC Best Local Similarity 75.0%; Pred. No. 59;  
CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
-----  
CC YQ 2 GVITXIRP 9  
CC 75 GVITQITP 82  
-----  
CC RESULT 10  
CC PIT1 MELGA STANDARD; PRT; 370 AA.  
CC ID PIT1 MELGA STANDARD; PRT; 370 AA.  
CC AC Q05749;  
CC DT 01-FEB-1995 (Rel. 31, Created)  
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Pituitary-specific positive transcription factor 1 (Pit-1) (Growth  
CC hormone factor 1) (GHP-1).  
CC GN POU1F1 OR PIT1.  
CC OS Meleagris gallopavo (Common turkey).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.  
CC OX NCBI\_TaxID=9103;  
CC [1]  
CC RN SEQUENCE FROM N.A.  
CC RP TISSUE=Pituitary;  
CC RX MEDLINE=93039671; PubMed=1418622;  
CC RA Wong E.A., Sileby J.L., el Halawani M.E.;  
CC RT "Complementary DNA cloning and expression of Pit-1/GHP-1 from the  
CC domestic turkey".  
CC RL DNA Cell Biol. 11:651-660 (1992).  
CC [2]  
CC RN REVISIONS, AND ALTERNATIVE SPLICING.  
CC RP Wong E.A., Sharova L., Kurima K., Weatherly K.L.;  
CC RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Transcription factor that activates growth hormone and  
CC prolactin genes. Specifically binds to the consensus sequence 5'-  
CC TAAAT-3'.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=PIT-1beta\*;  
CC ISOID=Q05749-1; Sequence=Displayed;  
CC Name=PIT-1;  
CC ISOID=Q05749-2; Sequence=VSP\_002318;  
CC Name=PIT-1\*;  
CC ISOID=Q05749-3; Sequence=VSP\_002319;  
CC TISSUE SPECIFICITY: Pituitary Gland.  
CC -!- SIMILARITY: Belongs to the POU transcription factor family.  
CC Class=1 subfamily.  
CC -!- SIMILARITY: Contains 1 homeobox domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U18928; AAB04690.1; -;  
CC EMBL; U62732; AAB04690.1; JOINED.  
CC EMBL; U18923; AAB04690.1; JOINED.  
CC EMBL; U18924; AAB04690.1; JOINED.  
CC EMBL; U18925; AAB04690.1; JOINED.  
CC EMBL; U18926; AAB04690.1; JOINED.  
CC EMBL; U18927; AAB04690.1; JOINED.  
CC EMBL; U18928; AAB04690.1; JOINED.  
CC EMBL; U18928; AAB04691.1; -;  
CC EMBL; U62732; AAB04691.1; JOINED.

**CC -!- FUNCTION:** Required for the insertion of integral membrane proteins

into the membrane. Probably plays an essential role in the integration of proteins of the respiratory chain complexes. Involved in integration of membrane proteins that insert dependently and independently of the Sec translocase complex (By similarity).

CC -1- SUBUNIT: Specifically interacts with transmembrane segments of nascent integral membrane proteins during membrane integration (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (By similarity).

CC -1- SIMILARITY: Belongs to the OXA1/oxaA family. Subfamily 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: BX572093; CAE19645.1; -

CC HAMAP: MF\_01810; -, 1.

CC TRANSMEM 20 42 POTENTIAL.

CC TRANSMEM 266 328 POTENTIAL.

CC TRANSMEM 303 325 POTENTIAL.

CC SEQUENCE 379 AA; 41262 MW; F5E9029078B36AF1 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 379;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
|::|::|  
DB 220 GLVTAIRP 227

RESULT 12

EF-1A AERPE STANDARD; PRT; 437 AA.

CC QSYAVQ;

CC 30-MAY-2000 (Rel. 39, Created)

CC DT 30-MAY-2000 (Rel. 39, Last sequence update)

CC DT 28-FEB-2003 (Rel. 41, Last annotation update)

CC DE Elongation factor 1-alpha (EF-1-alpha) (Elongation factor Tu) (EF-Tu).  
TU OF APE1844.

CC OS Aeropyrum pernix.

CC OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
CC OC Desulfurococaceae; Aeropyrum.

CC OX NCBI\_TaxID=56636;  
CC [1]

CC SEQUENCE FROM N. A.  
CC STRAIN=K1;  
CC MEDLINE=99310339; PubMed=10382966;  
CC RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
CC RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,  
CC RA Hoshiyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
CC RA Takamiya M., Masuda N., Funahashi T., Tanaka T., Kudoh Y.,  
CC RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,  
CC RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
CC RT "Complete genome sequence of an aerobic hyper-thermophilic  
CC crenarchaeon, Aeropyrum pernix K1.";  
CC RL DNA Res. 6:83-101(1999).

CC -1- FUNCTION: This protein promotes the GTP-dependent binding of  
CC aminoacyl-tRNA to the A-site of ribosomes during protein  
CC biosynthesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: Belongs to the GTP-binding elongation factor family.  
CC EF-Tu/EF-1A subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AP000062; BAA80848.1; -  
PIR; C72570; C72570.  
HSSP; P07157; LAIP.  
DR HAWAP; MF\_00118; -; 1.  
DR InterPro; IPR004539; EFl\_alpha.  
DR InterPro; IPR000795; EF\_GTPbind.  
DR InterPro; IPR004160; EFTU\_Cterm.  
DR InterPro; IPR004161; EFTU\_D2.  
DR InterPro; IPR009001; Elong\_init\_C.  
DR InterPro; IPR009000; translat\_factor.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
DR PRINTS; PR00315; ELONGATNFCT.  
DR TIGRFAMs; TIGR00483; EF-1\_alpha; 1.  
DR PROSITE; PS00301; EFATOR\_GTP; 1.  
KW Elongation factor; Protein biosynthesis; GTP-binding;  
FT NP\_BIND 13 20 GTP (BY SIMILARITY).  
FT NP\_BIND 90 94 GTP (BY SIMILARITY).  
FT NP\_BIND 152 155 GTP (BY SIMILARITY).  
FT SEQUENCE 437 AA; 48662 MW; D6AE87FEB8AD003C CRC64;

Query Match 75.7%; Score 28; DB 1; Length 437;  
Best Local Similarity 37.5%; Pred. No. 72;  
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| : : : : :  
DB 422 GIVTDVKP 429

RESULT 13  
Y4UC RHISN STANDARD; PRT; 512 AA.  
AC Q53197;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical aldehyde-dehydrogenase like protein Y4UC (EC 1.2.1.-).  
GN Y4UC.  
OS Rhizobium sp. (strain NGR234).  
OG Plasmid sym pNGR234a.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
OX NCBI\_TaxID=394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97305956; PubMed=9163424;  
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
RA Perret X.;  
RT "Molecular basis of symbiosis between Rhizobium and legumes.";  
RL Nature 387:394-401(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96389014; PubMed=8796346;  
RA Freiberg C., Perret X., Broughton W.J., Rosenthal A.;  
RT "Sequencing the 500-kb GC-rich symbiotic replicon of Rhizobium sp. NGR234 using dye terminators and a thermostable 'sequenase': a beginning.";  
RL Genome Res. 6:590-600(1996).  
CC -!- FUNCTION: COULD BE A SUCCINATE-SEMIALDEHYDE DEHYDROGENASE (NADP+).  
CC -!- PATHWAY: 4-aminobutyrate (GABA) degradation.  
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Z68203; CRA92404.1; -  
EMBL; AE000099; AAB91875.1; -  
HSSP; P51977; 1BXS.  
DR InterPro; IPR002086; Aldehyde\_dehydr.  
DR Pfam; PF00171; aldehyd; 1.  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; FALSE\_NEG.  
DR PROSITE; PS00687; ALDEHYDE\_DEHYDR\_GLU; 1.  
KW Hypothetical protein; Oxidoreductase; Plasmid.  
FT ACT\_SITE 286 286 BY SIMILARITY.  
FT ACT\_SITE 320 320 BY SIMILARITY.  
FT SEQUENCE 512 AA; 55360 MW; B8B862162FD251E9 CRC64;

Query Match 75.7%; Score 28; DB 1; Length 512;  
Best Local Similarity 62.5%; Pred. No. 83;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
| : : : : :  
DB 182 GIITAITP 189

RESULT 14  
ATSI MOUSE STANDARD; PRT; 968 AA.  
ID ATSI\_MOUSE  
AC P97857; OS4768;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
GN ADAMTS1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ;  
RX MEDLINE=98110583; PubMed=9441751;  
RA Kuno K., Lizaes H., Ohno S., Matsushima K.;  
RT "The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
RL Genomics 46:466-471(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97150761; PubMed=8995297;  
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
RA Matsushima K.;  
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene.";  
RL J. Biol. Chem. 272:556-562(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Limb and Mammary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udutin T.B., Teshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McSwan P.U., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.O., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[4]  
RN CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.  
RP MEDLINE=9303657; PubMed=10373500;  
RA Kuno K., Terashima Y., Matsushima K.;  
RT "ADAMTS-1 is an active metalloproteinase associated with the  
RT extracellular matrix";  
RL J. Biol. Chem. 274:18821-18826(1999).  
[5]  
RN FUNCTION.  
RP MEDLINE=20389569; PubMed=10930576;  
RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,  
RA Ohno H., Matsushima K.;  
RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
RL FEBS Lett. 478:241-245(2000).  
[6]  
RN FUNCTION, AND INDUCTION.  
RP MEDLINE=20243757; PubMed=10781075;  
RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
RA Richards J.S.;  
RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and  
RT cathepsin L proteases";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
CC involved in its turnover. Has angiogenic inhibitor activity (By  
CC similarity). Active metalloproteinase, which may be associated with  
CC various inflammatory processes as well as development of cancer  
CC cachexia. May play a critical role in follicular rupture (By  
CC similarity).  
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692  
CC site, within the chondroitin sulfate attachment domain.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix.  
CC -!- INDUCTION: Induced in vitro in colon adenocarcinoma cells by  
CC interleukin-1, or in vivo in kidney and heart by  
CC lipopolysaccharide. Also induced by LH stimulation in granulosa  
CC cells of preovulatory follicles.  
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important  
CC for a tight interaction with the extracellular matrix.  
CC -!- PTM: The precursor is cleaved by a furin endopeptidase.  
CC -!- SIMILARITY: Belongs to peptidase family M12B.  
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.  
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a  
CC frameshift in position 7.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC -----  
CC EMBL; AB001735; BAA24501.1; ALT INIT.  
CC EMBL; D67076; BAA11088.1; ALT\_FRAME.  
CC EMBL; BC040382; AAH40382.1; -;  
CC EMBL; BC050834; AAH50834.1; -;  
CC MEROPS; M12.222; -;  
CC MGD; MGI:109249; Adamts1.  
CC InterPro; IPR006586; ADAM cysteine.  
CC InterPro; IPR001762; Disintegrin.  
CC InterPro; IPR006025; Pept\_M\_Zn\_BS.  
CC InterPro; IPR001590; Peptidase\_M12B.  
CC InterPro; IPR002870; Peptidase\_M12B\_N.

DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP 1.  
DR Pfam; PF01562; Rep\_M12B\_Propag; 1.  
DR Pfam; PF01421; RepPolysin; 1.  
DR Pfam; PF00090; tsp\_1; 3.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00608; ACR; 1.  
DR SMART; SM00209; TSP1; 3.  
DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
DR PROSITE; PS50092; tsp\_1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
KW Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;  
KW Repeat; Extracellular matrix; Heparin-binding.  
FT SIGNAL 1 48  
FT PROPEP 49 253  
FT CHAIN 254 968  
FT DOMAIN 254 476  
FT DOMAIN 477 559  
FT DOMAIN 560 615  
FT DOMAIN 618 725  
FT DOMAIN 726 850  
FT DOMAIN 855 911  
FT DOMAIN 912 968  
FT DOMAIN 135 199  
FT SITE 206 206  
FT METAL 402 402  
FT ACT\_SITE 403 403  
FT METAL 406 406  
FT METAL 412 412  
FT CARBOHYD 548 548  
FT CARBOHYD 721 721  
FT CARBOHYD 765 765  
FT CARBOHYD 783 783  
FT CARBOHYD 946 946  
FT MUTAGEN 403 403  
FT CONFLICT 335 335  
FT CONFLICT 425 425  
FT SEQUENCE 968 AA; 105841 MW; 42EBDA55499FB6C1 CRC64;  
SQ  
Query Match 75.7%; Score 28; DB 1; Length 968;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVITXIRP 9  
Db 730 GIVTSTRP 737  
  
RESULT 15  
UBAL\_WHEAT  
ID UBAL\_WHEAT STANDARD; PRT; 1051 AA.  
AC P20973;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ubiquitin-activating enzyme E1.  
GN UB1  
OS Triticum aestivum (Wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
OC Triticeae; Triticum.  
OC NCBI\_TaxID=4565;  
[1]  
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RP STRAIN=cv. Augusta;  
RC MEDLINE=90368797; PubMed=2203788;  
RA Hatfield P.M., Callis J., Vierstra R.D.;  
RT "Cloning of ubiquitin activating enzyme from wheat and expression of  
RT a functional protein in Escherichia coli.";  
RL J. Biol. Chem. 265:15813-15817(1990).  
[2]

```

RP MUTAGENESIS OF CYSTEINE RESIDUES, AND ACTIVE SITE.
RX MEDLINE-92340519; PubMed-1634524;
RA Hatfield P.M., Vierstra R.D.;
RT "Multiple forms of ubiquitin-activating enzyme E1 from wheat.
RL Identification of an essential cysteine by in vitro mutagenesis.";
CC J. Biol. Chem. 267:14799-14803(1992).
CC -!- FUNCTION: Activates ubiquitin by first adenylating with ATP its
CC carboxy-terminal glycine residue and thereafter linking this
CC residue to the side chain of a cysteine residue in E1, yielding an
CC ubiquitin-E1 thiolester and free AMP.
CC -!- PATHWAY: Ubiquitin conjugation; first step.
CC -!- SUBUNIT: Monomer.
CC -!- PTM: The N-terminus is blocked.
CC -!- MISCELLANEOUS: There are two active sites within the E1 molecule,
CC allowing it to accommodate two ubiquitin moieties at a time, with a
CC new ubiquitin forming an adenylate intermediate as the previous
CC one is transferred to the thiol site.
CC -!- MISCELLANEOUS: There are multiple genes encoding E1 in wheat.
CC -!- SIMILARITY: Belongs to the ubiquitin-activating E1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M55604; AAA34308.1; -.
DR InterPro; IPR009036; Moeb.
DR InterPro; IPR000594; ThiP_domain.
DR InterPro; IPR000127; UBact_repeat.
DR InterPro; IPR000011; Uqtin-activ_enz.
DR Pfam; PF00899; ThiF; 2.
DR 2fam; PF02134; UBACT; 2.
DR TIGRFAMs; TIGR01408; UBcl; 1.
DR PROSITE; PS00536; UBIQUITIN_ACTIVAT_1; 1.
DR PROSITE; PS00865; UBIQUITIN_ACTIVAT_2; 1.
DR KX ubl conjugation pathway; Ligase; Multigene family; Repeat.
FT ACT_SITE 626 626
FT DOMAIN 56 605 2 APPROXIMATE REPEATS.
FT REPEAT 56 194 1-1.
FT REPEAT 453 605 1-2.
FT SEQUENCE 1051 AA; 117007 MW; 9BAB85FE3BACA621 CRC64;
Query Match 75.7%; Score 28; DB 1; Length 1051;
Best Local Similarity 37.5%; Pred. NO. 1.6e+02;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 2 GVITKRP 9
DB 286 GIVTQVKP 293

```

Search completed: April 1, 2004, 17:39:24  
Job time : 7.52632 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 28.4211 Seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-2

Perfect score: 37

Sequence: 1 XGVITXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	91.9	407	16	Q8F364
2	33	89.2	139	2	Q84CK2
3	32	86.5	115	12	Q39919
4	32	86.5	376	16	Q82PR2
5	32	86.5	496	16	Q9A5P7
6	32	86.5	1472	16	Q9L2I4
7	31	83.8	59	9	Q859A7
8	31	83.8	166	16	Q07538
9	31	83.8	254	10	Q9LI08
10	31	83.8	269	2	Q9WWM2
11	31	83.8	288	16	Q8YXD0
12	31	83.8	289	16	Q8DIJ3
13	31	83.8	318	11	Q8VFT1
14	31	83.8	344	9	Q857L6
15	31	83.8	486	5	Q26617
16	31	83.8	498	16	Q9KQ09

Q82X57 nitrosomona  
Q75W3 brachydanio  
Q7ZY6 brachydanio  
Q7ZY8 brachydanio  
Q75W2 brachydanio  
Q82Y7 yersinia pe  
Q8NJ75 agaricus bi  
Q8KJ22 mus musculu  
Q8KQV3 vibrio chol  
Q97DA0 clostridium  
Q8RXX2 providencia  
Q7ZY04 xenopus lae  
Q8H62 oryza sativ  
Q98JP7 rhizobium 1  
Q8UJ4 bacillus an  
Q8H62 bacillus ce  
Q94D96 oryza sativ  
Q90Z37 brachydanio  
Q8DA47 vibrio vuln  
Q8S08 oryza sativ  
Q8E746 streptococ  
Q8E1N0 streptococ  
Q932P1 staphylococ  
Q7X6K7 oryza sativ  
Q9FZy6 streptomyce  
Q8KUI3 actinosynne  
Q9A7C0 caulobacte  
Q8NTV2 corynebacte  
Q8FU44 corynebacte

#### ALIGNMENTS

RESULT 1

Q8F364  
ID Q8F364 PRELIMINARY; PRT; 407 AA.  
AC Q8F364;  
DT 01-VAR-2003 (TREMBLrel. 23, Created)  
DT 01-VAR-2003 (TREMBLrel. 23, Last sequence update)  
DE Capb protein, Mur ligase family.  
GN CAPB OR LA2547.  
OS Leptospira interrogans.  
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.  
OX NCBI\_TaxID=173;  
RN [1]\_TaxID=173;  
RP SEQUENCE FROM N.A.  
RC STRAIN=56E01 / Serogroup Icterohaemorrhagiae / Serovar lai;  
RA Ren S.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE011422; AAN49746.1;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0003677; F:DNA binding; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0045227; F:capsule polysaccharide biosynthesis; IEA.  
DR InterPro; IPR008337; CapB.  
DR InterPro; IPR001005; MYB DNA binding.  
DR PRINTS; PR01758; CAPSULEPROTB.  
DR PROSITE; PS00037; MYB.1; 1.  
KW Ligase; Complete proteome.  
SQ SEQUENCE 407 AA; 46759 MW; APL2FAB606746BAB CRC64;

Query Match 91.9%; Score 34; DB 16; Length 407;  
Best Local Similarity 87.5%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 QVITXIRP 9

Db 132 QVITXIRP 139



## RESULT 2

Q84CK2 Q84CK2 PRELIMINARY; PRT; 139 AA.  
 AC Q84CK2;  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Kinul.  
 GN KINUL.  
 OS Streptomyces murayamaensis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=224537;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yin X.H., Mahadevan B., Grochowski L., Proteau P.J.;  
 RT "Molecular cloning and sequence of the kinamycin angucycline type II  
 polyketide synthase gene cluster from Streptomyces murayamaensis";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY228175; AAC65341.1;  
 DR InterPro; IPR009002; FMN binding.  
 SQ SEQUENCE 139 AA; 148.9 MW; 65A27123A80349E4 CRC64;

Query Match 89.2%; Score 33; DB 2; Length 139;  
 Best Local Similarity 75.0%; Pred. No. 12;  
 Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 GVITXIRP 9  
 DB 28 GVLTIRP 35

## RESULT 3

O39919 O39919 PRELIMINARY; PRT; 115 AA.  
 AC O39919;  
 DT 01-JAN-1998 (TREMELrel. 05, Created)  
 DT 01-JAN-1999 (TREMELrel. 09, Last sequence update)  
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
 DE UL41 protein (Fragment).  
 GN UL41.  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Dargan D.J., Jamieson F.S., Maclean J., Dolan A., Addison C.,  
 RT "The published DNA sequence of human cytomegalovirus strain AD169  
 lacks 929 base pairs of DNA affecting genes UL42 and UL43.";  
 RL J. Virol. 71:9833-9836(1997).  
 RN [2]

QY 2 GVITXIRP 9  
 DB 28 GVLTIRP 35

Query Match 86.5%; Score 32; DB 12; Length 115;  
 Best Local Similarity 62.5%; Pred. No. 17;  
 Matches 5; Conservative 2; Mismatches 1; Indels 1; Gaps 0;

QY 2 GVITXIRP 9

Db 8 GIITLRP 15

## RESULT 4

Q82PR2 Q82PR2 PRELIMINARY; PRT; 376 AA.  
 AC Q82PR2;  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Putative dehydrogenase.  
 GN SAV810.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=21477403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Rattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 avermitilis: deducing the ability of producing secondary  
 metabolites";  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Rattori M., Omura S.;  
 RT "Complete genome sequence and comparative analysis of the industrial  
 microorganism Streptomyces avermitilis";  
 RL Nat. Biotechnol. 21:526-531(2003).  
 DR EMBL; AP005024; BAC68520.1;  
 DR GO; GO:0015036; F:disulfide oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001327; FAD\_pyr\_redox.  
 DR Pfam; PF00070; Pyr\_redox; 1.  
 DR PRINTS; PR00368; FADPNR.  
 KW Complete Proteome.  
 SQ SEQUENCE 376 AA; 39115 MW; CF70B8B31BBB8D2C CRC64;

Query Match 86.5%; Score 32; DB 16; Length 376;  
 Best Local Similarity 62.5%; Pred. No. 55;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9  
 DB 181 GVTEVRP 188

## RESULT 5

O9A5P7 O9A5P7 PRELIMINARY; PRT; 496 AA.  
 AC O9A5P7;  
 DT 01-JUN-2001 (TREMELrel. 17, Created)  
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE Acid-CoA ligase, putative.  
 GN CC2400.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

```

RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Pocock I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Smolava M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl Acad. Sci U.S.A. 98:4136-4141(2001).
DR ENBL, AB005909, AK24371.1; -.
DR PIR, G87546; G87546.
DR TIGR, CC2400; -.
DR GO, GO:0016874; F:ligase activity; IEA.
DR GO, GO:0008152; P:metabolism; IEA.
DR InterPro, IPR000873; AMP-bind.
DR Pfam, PF00501; AMP-binding; 1.
DR PROSITE, PS00455; AMP BINDING; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 495 AA; 53547 MW; 81A26FD47947F37D CRC64;

Query Match 86.5%; Score 32; DB 16; Length 496;
Best Local Similarity 75.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 196 GLITSIRP 203

RESULT 6
Q9L2I4 PRELIMINARY; PRT; 1472 AA.
AC Q9L2I4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transport ATPase.
DE SC02499 OR SC121.02C.
GN Streptomyces coelicolor.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (JAN-2000) to the ENBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Cerdano A.M., Parkhill J., Barrell B.G., Denapante D., Eichner A., Cullum J.,
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
EX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;

```

```

RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR ENBL, AL399112; CAB69720.1; -.
DR GO, GO:0016020; C:membrane; IEA.
DR GO, GO:0005524; F:ATP binding; IEA.
DR GO, GO:0015662; F:ATPase activity; IEA.
DR GO, GO:0016787; F:hydrolase activity; IEA.
DR GO, GO:0006812; P:cation transport; IEA.
DR GO, GO:0008152; P:metabolism; IEA.
DR InterPro, IPR001757; ATPase E1-E2.
DR InterPro, IPR008250; E1-E2 ATPase_reg.
DR InterPro, IPR005834; Hydrolase.
DR Pfam, PF00122; E1-E2 ATPase; 1.
DR Pfam, PF00702; Hydrolase; 1.
DR PRINTS, PR00119; CATAPASE.
DR TIGRfams, TIGR01494; ATPase P-type; 5.
DR PROSITE, PS00154; ATPase_E1-E2; 1.
KW Complete proteome.
SQ SEQUENCE 1472 AA; 151649 MW; A71389768FE4D28A CRC64;

Query Match 86.5%; Score 32; DB 16; Length 1472;
Best Local Similarity 62.5%; Pred. No. 2.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 466 GVITVVRP 473

RESULT 7
Q859A7 PRELIMINARY; PRT; 59 AA.
AC Q859A7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gp51.
OS Bacteriophage phi-Btl.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX NCBI_TaxID=225588;
RN [1]
RP SEQUENCE FROM N.A.
RC Gregory M.A.;
RA Gregory M.A.;
RL Submitted (MAR-2003) to the ENBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Gregory M.A.;
RT "Characterisation and Evolution of Homimmune Streptomyces
RT Bacteriophages.";
RL Thesis (2000), Department of Genetics, University of Nottingham,
RL Nottingham, UNITED KINGDOM.
DR ENBL, AJ550940; CAD80118.1; -.
SQ SEQUENCE 59 AA; 5827 MW; 7D7F7B9F8CDC7020 CRC64;

Query Match 83.8%; Score 31; DB 9; Length 59;
Best Local Similarity 62.5%; Pred. No. 15;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 35 GVITAVKP 42

RESULT 8
Q07538 PRELIMINARY; PRT; 166 AA.
AC Q07538;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yhaw.

```

GN YHAW.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriss R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Broillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kaashara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Maue J., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara K., Oudega B., Park S.H.,  
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle D., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takouchi M., Tamakoshi A., Tanaka T., Terpatra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zungstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis";  
RL Nature 390:249-256(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Y14080; CAA74454.1; -;  
DR EMBL; Z99109; CAB12823.1; -;  
DR PIR; A69820; A69820.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 166 AA; 19025 MW; 0EB8578CD64F1D63 CRC64;

Query Match 83.8%; Score 31; DB 16; Length 166;  
Best Local Similarity 62.5%; Pred. No. 42;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXRP 9  
Db 108 GVITXRP 115  
||:|:|:|  
RESULT 9  
ID Q9LI08 PRELIMINARY; PRT; 254 AA.  
AC Q9LI08;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein.  
OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
RT clone: P0029D06.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP001552; BAA93016.1; -;  
DR Gramene; O9LI08; -;  
KW Hypothetical protein.  
SQ SEQUENCE 254 AA; 26867 MW; 65366CEB342D8192 CRC64;  
Query Match 83.8%; Score 31; DB 10; Length 254;  
Best Local Similarity 62.5%; Pred. No. 63;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXRP 9  
Db 140 GVITXRP 147  
||:|:|:|  
RESULT 10  
ID Q9WWM2 PRELIMINARY; PRT; 269 AA.  
AC Q9WWM2;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Synecococcus PCC6301 idIA gene and ORF1 to 7 (fragment).  
OS Synecococcus sp. (strain PCC 6301) (Anacystis nidulans).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
OX NCBI\_TaxID=1139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96425891; PubMed=8828233;  
RX Michel K., Thole H.H., Pistorius E.;  
RT "IdIA, a 34 kDa protein in the cyanobacteria Synecococcus sp. strains  
RT PCC 6301 and PCC 7942, is required for growth under iron and manganese  
RT limitations.";  
RL Microbiology 142:2635-2645(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=98262539; PubMed=9599805;  
RX Michel K.P., Exss-Sonne P., Scholten-Beck G., Kahmann U., Ruppel H.G.,  
RA Pistorius E.K.;  
RT "Immunocytochemical localization of IdIA, a protein expressed under  
RT iron or manganese limitation in the mesophilic cyanobacterium  
RT Synecococcus PCC 6301 and the thermophilic cyanobacterium  
RT Synecococcus elongatus.";  
RL Planta 205:73-81(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=9937081; PubMed=10411274;  
RX Michel K.P., Krueger F., Puhler A., Pistorius E.K.;  
RA "Molecular characterization of idIA and adjacent genes in the  
RA cyanobacteria Synecococcus sp. strains PCC 6301 and PCC 7942.";  
RL Microbiology 145:1473-1484(1999).  
DR EMBL; Z48754; CAB45539.1; -;  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR001851; Bac\_inmem\_transp.  
DR Pfam; PF02653; BPD\_transp\_2; 1.  
FT NON TER 269  
SQ SEQUENCE 269 AA; 28391 MW; 645B961A6AE9867D CRC64;

Query Match 83.8%; Score 31; DB 2; Length 269;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

DR GO; GO:0005215; F-transporter activity; IEA.
DR GO; GO:0006810; P-transport; IEA.
DR InterPro; IPR001851; Bac_innem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete proteome.
SQ SEQUENCE 289 AA; 30623 MW; CF73BD3BB884BC19 CRC64;

Query Match      83.8%; Score 31; DB 16; Length 289;
Best Local Similarity 52.5%; Pred.No. 72;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
   |:|:|
DB 209 GLITAVRP 216

RESULT 13
QSVFT1 PRELIMINARY; PET; 318 AA.
ID QSVFT1
AC QSVFT1;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Olfactory receptor MOR264-6 (Olfactory receptor
DB GA_xK60272Q125-48676316-48677272).
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang X., Firestein S.J.;
RT "The olfactory receptor gene superfamily of the mouse.";
RL Nat. Neurosci. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Adams M.;
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Young J.M., Shykind B.M., Lane R.P., Tonnes-Priddy L., Ross J.A.,
RA Walker M., Williams E.M., Axel R., Trask B.J.;
RT "odorant receptor ESTs demonstrate olfactory expression of over 400
RL genes, extensive alternate splicing, and unequal expression levels.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Sanders K.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY073435; AAL61098.1; -.
DR EMBL; AY318309; AAF71535.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 318 AA; 35872 MW; 4CAF8AB634985779 CRC64;

Query Match      83.8%; Score 31; DB 11; Length 318;
Best Local Similarity 62.5%; Pred.No. 79;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVITXIRP 9
   |:|:|
DB 262 GLITLVRP 269

RESULT 14

```

```

Q857L6
ID Q857L6 PRELIMINARY; PRT; 344 AA.
AC Q857L6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Gp4.
OS Mycobacteriophage Bx22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=205870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22592660; PubMed=12705866;
RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,
RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Fannunzio N.R.,
RA Brucker W., Kumar V., Kandassamy J., Keenan L., Bardarov S.,
RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,
RA Hatfull G.F.;
RT "Origins of highly mosaic mycobacteriophage genomes.";
RL Cell 113:171-182(2003).
DR EMBL; AY129332; AAN01760.1; -.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 3.
SQ SEQUENCE 344 AA; 33535 MW; C48AGD28185A5989 CRC64;

Query Match 83.8%; Score 31; DB 9; Length 344;
Best Local Similarity 62.5%; Pred. No. 86;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 43 GVLTRVP 50

RESULT 15
ID Q26617 PRELIMINARY; PRT; 486 AA.
AC Q26617;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE SpGCF1.
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95301110; PubMed=7781310;
RA Zeller R.W., Coffman J.A., Harrington M.G., Britten R.J.,
RA Davidson E.H.;
RT "SpGCF1, a sea urchin embryo DNA-binding protein, exists as five
RT nested variants encoded by a single mRNA.";
RL Dev. Biol. 169:713-727(1995).
DR EMBL; U18784; AA86479.1; -.
DR InterPro; IPR006162; Peptide S.
DR PROSITE; PS00012; PHOSPHOPANTHETINE; 1.
SQ SEQUENCE 486 AA; 54616 MW; A9ABF23F826357 CRC64;

Query Match 83.8%; Score 31; DB 5; Length 486;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVITXIRP 9
Db 26 GVITDIP 33

Search completed: April 1, 2004, 17:38:37
Job time : 30.4211 secs

```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 43.1053 Seconds  
(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-3

Perfect score: 39

Sequence: 1 XGVGTWIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	89.7	193	2	AAY74014 Human pro
2	31	79.5	229	2	Aaw70263 Amino aci
3	31	79.5	229	6	ABU01701 S. pneumo
4	31	79.5	281	6	ABU29473 Protein e
5	31	79.5	342	5	ABG91685 Purine/py
6	31	79.5	353	5	ABG91498 Purine/py
7	31	79.5	361	5	ABG91515 Purine/py
8	31	79.5	729	6	ADA36749 Acinetoba
9	30	76.9	8	7	ADD31284 Angiogene
10	30	76.9	8	7	ADD31284 Angiogene
11	30	76.9	9	4	AAU15042 Antiangio
12	30	76.9	10	4	AAU14987 Antiangio
13	30	76.9	12	3	AB35743 Glucodext
14	30	76.9	64	4	AAU43108 Propionib
15	30	76.9	64	4	AAU43108 Propionib
16	30	76.9	123	7	ADC87823 Ribosomal
17	30	76.9	241	2	AAW36812 Helicobac
18	30	76.9	246	5	ABG91662 Purine/py
19	30	76.9	259	6	ABU42006 Protein e
20	30	76.9	273	5	ABG93435 Herbicida
21	30	76.9	328	5	ABG91542 Purine/py
22	30	76.9	335	2	AAK89959 FMDV lb p
23	30	76.9	352	5	ABG91503 Purine/py
24	30	76.9	355	4	AAU04848 Micronono
25	30	76.9	355	5	ABG91499 Purine/py

26	30	76.9	355	5	ABG91505	Abg91505 Purine/py
27	30	76.9	355	5	ABG91500	Abg91500 Purine/py
28	30	76.9	355	6	ABP99275	Abp99275 Orthosomy
29	30	76.9	355	6	ABP99273	Abp99273 Orthosomy
30	30	76.9	355	6	ABP76704	Abp76704 Streptom
31	30	76.9	385	5	ABG91607	Abg91607 Purine/py
32	30	76.9	424	2	AAV74075	AAV74075 Human pro
33	30	76.9	463	6	ABU34851	Abu34851 Protein e
34	30	76.9	463	6	ABU36436	Abu36436 Protein e
35	30	76.9	470	5	ABB47522	Abb47522 Listeria
36	30	76.9	533	6	ABU33860	Abu33860 Protein e
37	30	76.9	536	6	ABU33716	Abu33716 Protein e
38	30	76.9	843	3	ABG91630	Abg91630 Purine/py
39	30	76.9	1047	3	AA35740	AA35740 Glucodext
40	30	76.9	2872	6	ABU32213	Abu32213 Protein e
41	30	76.9	19938	6	ABB98398	Abb98398 Streptom
42	30	76.9	19938	6	ABP76678	Abp76678 Streptom
43	29	74.4	8	7	ADD31265	Add31265 Angiogene
44	29	74.4	8	7	ADD31294	Add31294 Angiogene
45	29	74.4	8	7	ADE83629	Ade83629 Antiangi

## ALIGNMENTS

RESULT 1  
AAY74014  
ID AAY74014 standard; protein; 193 AA.  
XX  
AC AAY74014;  
XX  
DT 14-MAR-2000 (first entry)  
XX  
DE Human prostate tumor EST fragment derived protein #201.  
XX  
KW Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;  
KW treatment.  
XX  
OS Homo sapiens.  
XX  
PN DE19820190-A1.  
XX  
PD 04-NOV-1999.  
XX  
PF 28-APR-1998; 98DE-01020190.  
XX  
PR 28-APR-1998; 98DE-01020190.  
XX  
(META-) METAGEN GES GENOMFORSCHUNG MBH.  
PA Rosenthal A, Specht T, Hinzmann B, Schmitt A, Piliarsky C, Dahl E;  
XX  
XX WPI; 1999-621386/54.  
DR N-PSDB; AA252924.  
XX  
XX New human nucleic acid sequences from pancreatic tumors, and related  
PT proteins.  
PT  
PS Claim 23; Page 393; 502pp; German.  
XX  
XX This invention describes novel polypeptides and their encoding nucleic  
CC acids derived from human pancreatic tumor tissue which have cytostatic  
CC activity. The sequences are also useful in producing pharmaceutical  
CC compositions for treatment of pancreatic tumors. AAY73814-Y74252  
CC represent protein fragments encoded by the human pancreatic tumor cDNA  
CC library derived expressed sequence tag (EST) sequences represented in  
CC AA252859-253014  
XX  
XX  
SQ Sequence 193 AA;

Query Match 89.7%; Score 35; DB 2; Length 193;  
Best Local Similarity 75.0%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY      2 GVGTXIRP 9
DB      87 GVGTSLRP 94

RESULT 2
AAW70263
ID      AAW70263 standard; protein; 229 AA.
XX      AC      AAW70263;
XX      DT      20-NOV-1998 (first entry)
XX      DE      Amino acid sequence of Streptococcus pneumoniae licC protein.
XX      KW      Choline kinase; licC gene; antagonist; antibody; T cell; immune response;
XX      KW      vaccination; otitis media; conjunctivitis; pneumonia; bacteraemia;
XX      KW      pleural empyaema; meningitis; sinusitis; endocarditis.
XX      OS      Streptococcus pneumoniae.
XX      FH      Key Location/Qualifiers
FT      Misc-difference 1 /note= "encoded by GTG"
XX      PN      EP861890-A1.
XX      PD      02-SEP-1998.
XX      PF      27-FEB-1998; 98EP-00301482.
XX      PR      28-FEB-1997; 97US-0039210P.
XX      PA      (SMIK ) SMITHKLINE BEECHAM CORP.
XX      PI      Lonetto MA;
XX      DR      WPI; 1998-449107/39.
XX      DR      N-PSDB; AAV33009.
XX      PT      New DNA encoding Streptococcus pneumoniae licC polypeptides - used to
XX      PT      prevent, treat, diagnose and vaccinate against e.g. otitis media,
XX      PT      conjunctivitis, pneumonia, bacteraemia, pleural empyaema, meningitis,
XX      PT      sinusitis and endocarditis.
XX      PS      Claim 1; Page 4; 29pp; English.
XX      CC      This is the amino acid sequence of the Streptococcus pneumoniae licC
XX      CC      protein, a choline kinase used in the method of the invention. The licC
XX      CC      polypeptide is used for the treatment of an individual in need of licC
XX      CC      protein and the antagonist is used for the treatment of an individual
XX      CC      with the need to inhibit licC polypeptide. Inducing an immunological
XX      CC      response in a mammal comprises inoculating the mammal with the licC
XX      CC      polypeptide to produce antibody and/or T cell immune responses. LicC is
XX      CC      used to prevent, treat, diagnose and vaccinate against e.g. otitis media,
XX      CC      conjunctivitis, pneumonia, bacteraemia, pleural empyaema, meningitis,
XX      CC      sinusitis and endocarditis
XX      SQ      Sequence 229 AA;
XX      Query Match 79.5%; Score 31; DB 2; Length 229;
XX      Best Local Similarity 62.5%; Pred. No. 2.1e-02;
XX      Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
DB      9 GLGTLRP 16

RESULT 3
ABU01701
ID      ABU01701 standard; protein; 229 AA.
XX      AC      ABU01701;
XX      DT      23-OCT-2003 (revised)
XX      DT      11-FEB-2003 (first entry)
XX      DE      S. pneumoniae type 4 strain protein from coding region #1277.
XX      KW      Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
XX      KW      antiinflammatory; antibacterial; immunostimulant; auditory; respiratory;
XX      KW      gene therapy; vaccine.
XX      OS      Streptococcus pneumoniae; type 4 strain.
XX      PN      WO200277021-A2.
XX      PD      03-OCT-2002.
XX      PF      27-MAR-2002; 2002WO-IB002163.
XX      PR      27-MAR-2001; 2001GB-00007658.
XX      PA      (CHIR-) CHIRON SPA.
XX      PA      (GENO-) INST GENOMIC RES.
XX      PI      Masignani V, Tettelin H, Fraser C;
XX      DR      WPI; 2003-040579/03.
XX      DR      N-PSDB; ABX06989.
XX      PT      New proteins and nucleic acid molecules from Streptococcus pneumoniae,
XX      PT      useful as medicaments for treating or preventing a disease or infection
XX      PT      due to streptococcus bacteria, such as pneumonia, sepsis, otitis media or
XX      PT      ear infection.
XX      PS      Claim 1; SEQ ID NO 2554; 56pp; English.
XX      CC      The invention relates to a protein comprising or having at least 50%
XX      CC      identity to any of the 2469 amino acid sequences, identified in the
XX      CC      specification (available on a computer readable format), or its fragment,
XX      CC      expressed from 2469 of 2489 identified DNA coding regions from the
XX      CC      Streptococcus pneumoniae type 4 strain genomic sequence appearing as
XX      CC      ABS56454. Also included are an antibody which binds one of the proteins,
XX      CC      treating a patient by administering the protein, DNA or antibody (in a
XX      CC      composition), a kit comprising first and second primers, which are the
XX      CC      nucleic acid cited above or fragments between nucleotides 8-100 of a
XX      CC      sequence not defined in the specification, for amplifying a target
XX      CC      sequence contained within a Streptococcus nucleic acid sequence, where
XX      CC      the first primer is substantially complementary to the target sequence of
XX      CC      and the second primer is substantially complementary to the complement of
XX      CC      the target sequence, and where the parts of the primers having
XX      CC      substantial complementarity define the termini of the target sequence to
XX      CC      be amplified, assay comprising contacting a test compound with the
XX      CC      protein, and determining whether the test compound binds to the protein
XX      CC      and a Streptococcus pneumoniae bacterium, where one or more genes
XX      CC      encoding the proteins has been rendered inactive. The proteins, nucleic
XX      CC      acid molecules, antibody and compositions are useful as medicaments for
XX      CC      treating or preventing a disease or infection due to streptococcus
XX      CC      bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
XX      CC      media or ear infection. They are also useful in developing vaccines,
XX      CC      diagnostics and antibiotics. The methods are useful for identifying
XX      CC      immunodominant proteins. The present sequence is one of the 2469 proteins
XX      CC      expressed by the identified coding regions from the genomic sequence.
XX      CC      Note: The sequence data for this patent did not form part of the printed
XX      CC      specification, but was obtained in electronic format directly from WIPO
XX      CC      at ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
XX      CC      standardise OS field)
XX      SQ      Sequence 229 AA;
XX      Query Match 79.5%; Score 31; DB 6; Length 229;
XX      Best Local Similarity 62.5%; Pred. No. 2.1e-02;
XX      Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 GVGTXIRP 9  
|.:||.:||  
Db 9 GLGTRLRP 16

RESULT 4  
ABU29473  
ID ABU29473 standard; protein; 281 AA.  
XX  
AC ABU29473;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #15000.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Enterococcus faecalis.  
XX  
FN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PF 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.  
PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;  
PI  
DR N-PSDB; ACA33343.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 57397; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC on a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 281 AA;  
Query Match 79.5%; Score 31; DB 6; Length 281;  
Best Local Similarity 75.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|.:||.:||  
Db 59 GVGTYIRP 66

RESULT 5  
ABG91685  
ID ABG91685 standard; protein; 342 AA.  
XX  
AC ABG91685;  
XX  
DT 18-NOV-2002 (first entry)  
XX  
DE Purine/pyrimidine triphosphate type nucleotidyltransferase #270.  
XX  
KW Nucleotidyltransferase; enzyme; active site engineering;  
KW alpha-D-glucopyranosyl phosphate thymidyltransferase; Ep;  
KW substrate specificity; nucleotide sugar;  
KW glycosylated bioactive natural product.  
XX  
OS Myxococcus xanthus.  
XX  
FN WO200248331-A2.  
XX  
PD 20-JUN-2002.  
XX  
PF 13-DEC-2001; 2001WO-US047953.  
XX  
PR 13-DEC-2000; 2000US-0254927P.  
XX  
PA (SLOK) SLOAN KETTERING INST CANCER RES.  
XX  
PI Thorson JS, Nikilov DB;  
XX  
DR WPI; 2002-608282/65.  
XX  
CC Nucleotidyltransferase mutated at one or more amino acids, useful in  
CC the synthesis of nucleotide sugars.  
XX  
PS Claim 3; Page; 182pp; English.  
XX  
CC The invention relates to a Nucleotidyltransferase mutated at one or  
CC more amino acids selected from V173, G147, W224, N112, G175, D111, E162,  
CC T201, I200, E199, R195, L89, L89T, L109, Y146 or Y177 (with reference to  
CC the *Salmonella enterica* rmlA-encoded alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase, Ep, enzyme appearing as ABG91798). The mutations  
CC alter the substrate specificity of the enzymes. The mutants and methods  
CC involving them are used in the synthesis of nucleotide sugars for  
CC altering nucleotidyltransferase substrate specificity. The  
CC nucleotidyltransferase exhibits different substrate specificity for  
CC GTP, CTP, TTP, UTP and ATP than a non-mutated nucleotidyltransferase.  
CC The mutant may also exhibit a high degree of sequence identity to  
CC *Salmonella enterica* LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep) and can convert a wide variety of phosphates.  
CC The mutants can be exploited in the biosynthesis of glycosylated  
CC bioactive natural products of pharmacological use. The present sequence  
CC is a nucleotidyltransferase exhibiting a high degree of sequence  
CC identity to *Salmonella enterica* LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep). Note: The present sequence is not displayed  
CC in the specification but was obtained from Genbank  
XX  
SQ Sequence 342 AA;



Query Match 79.5%; Score 31; DB 5; Length 342;  
Best Local Similarity 62.5%; Pred. No. 3e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 9 GLGTRLRP 16

RESULT 6  
ABG91498  
ID ABG91498 standard; protein; 353 AA.  
XX  
AC ABG91498;  
XX  
DT 18-NOV-2002 (first entry)  
XX  
DE Purine/pyrimidine triphosphate type nucleotidyltransferase #83.

XX Nucleotidyltransferase; enzyme; active site engineering;  
KW alpha-D-glucopyranosyl phosphate thymidyltransferase; Ep;  
KW substrate specificity; nucleotide sugar;  
KW glycosylated bioactive natural product.

XX Streptomyces violaceoruber.

XX WO200248331-A2.

XX 20-JUN-2002.

XX 13-DEC-2001; 2001WO-US047953.

XX 13-DEC-2000; 2000US-0254927P.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX Thorson JS, Nikilov DB;

XX WPI; 2002-608282/65.

XX Nucleotidyltransferase mutated at one or more amino acids, useful in  
PT the synthesis of nucleotide sugars.

XX Claim 3; Page; 182pp; English.

XX The invention relates to a Nucleotidyltransferase mutated at one or  
CC more amino acids selected from V173, G147, W224, N112, G175, D111, E162,  
CC T201, I200, E199, R195, L89, L89T, L109, Y146 or Y177 (with reference to  
CC the Salmonella enterica rmlA-encoded alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase, Ep, enzyme appearing as ABG91798). The mutations  
CC alter the substrate specificity of the enzymes. The mutants and methods  
CC involving them are used in the synthesis of nucleotide sugars for  
CC altering nucleotidyltransferase substrate specificity. The  
CC nucleotidyltransferase exhibits different substrate specificity for  
CC GTP, CTP, TTP, UTP and ATP than a non-mutated nucleotidyltransferase.  
CC The mutant may also exhibit a high degree of sequence identity to  
CC Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep) and can convert a wide variety of phosphates.  
CC Bioactive natural products of pharmacological use. The present sequence  
CC is a nucleotidyltransferase exhibiting a high degree of sequence  
CC identity to Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep). Note: The present sequence is not displayed  
CC in the specification but was obtained from Genbank

XX Sequence 353 AA;

Query Match 79.5%; Score 31; DB 5; Length 353;  
Best Local Similarity 62.5%; Pred. No. 3.1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9

Db 9 GMGTRLRP 16

RESULT 7  
AEG91515  
ID AEG91515 standard; protein; 361 AA.

XX AEG91515;

XX 18-NOV-2002 (first entry)

XX Purine/pyrimidine triphosphate type nucleotidyltransferase #100.

XX Nucleotidyltransferase; enzyme; active site engineering;  
KW alpha-D-glucopyranosyl phosphate thymidyltransferase; Ep;  
KW substrate specificity; nucleotide sugar;  
KW glycosylated bioactive natural product.

XX Deinococcus radiodurans.

XX WO200248331-A2.

XX 20-JUN-2002.

XX 13-DEC-2001; 2001WO-US047953.

XX 13-DEC-2000; 2000US-0254927P.

XX (SLOK ) SLOAN KETTERING INST CANCER RES.

XX Thorson JS, Nikilov DB;

XX WPI; 2002-608282/65.

XX Nucleotidyltransferase mutated at one or more amino acids, useful in  
PT the synthesis of nucleotide sugars.

XX Claim 3; Page; 182pp; English.

XX The invention relates to a Nucleotidyltransferase mutated at one or  
CC more amino acids selected from V173, G147, W224, N112, G175, D111, E162,  
CC T201, I200, E199, R195, L89, L89T, L109, Y146 or Y177 (with reference to  
CC the Salmonella enterica rmlA-encoded alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase, Ep, enzyme appearing as ABG91798). The mutations  
CC alter the substrate specificity of the enzymes. The mutants and methods  
CC involving them are used in the synthesis of nucleotide sugars for  
CC altering nucleotidyltransferase substrate specificity. The  
CC nucleotidyltransferase exhibits different substrate specificity for  
CC GTP, CTP, TTP, UTP and ATP than a non-mutated nucleotidyltransferase.  
CC The mutant may also exhibit a high degree of sequence identity to  
CC Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep) and can convert a wide variety of phosphates.  
CC Bioactive natural products of pharmacological use. The present sequence  
CC is a nucleotidyltransferase exhibiting a high degree of sequence  
CC identity to Salmonella enterica LT2 alpha-D-glucopyranosyl phosphate  
CC thymidyltransferase (Ep). Note: The present sequence is not displayed  
CC in the specification but was obtained from Genbank

XX Sequence 361 AA;

Query Match 79.5%; Score 31; DB 5; Length 361;  
Best Local Similarity 62.5%; Pred. No. 3.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 18 GLGTRLRP 25

RESULT 8  
ADA36749

ID ADA36749 standard; protein; 729 AA.  
 XX AC ADA36749;  
 XX DT 20-NOV-2003 (first entry)  
 XX DE Acinetobacter baumannii protein #3910.  
 XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
 XX KW plant biocontrol agent.  
 XX OS Acinetobacter baumannii.  
 XX PN US6562958-B1.  
 XX PD 13-MAY-2003.  
 XX PF 04-JUN-1999; 99US-00328352.  
 XX PR 09-JUN-1998; 98US-0088701P.  
 XX PA (GENO-) GENOME THERAPEUTICS CORP.  
 XX PI Breton G, Bush D;  
 XX DR WPI; 2003-576092/54.  
 XX DR N-PSDB; ADA32623.  
 XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
 PT for diagnosing a bacterial disease, as components of antibacterial  
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
 PT plants.  
 XX Example; SEQ ID NO 8036; 328pp; English.  
 PS CC The invention relates to isolated Acinetobacter baumannii nucleic acids.  
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
 CC for diagnosing a bacterial disease, as components of antibacterial  
 CC vaccines, as targets for antibacterial drugs, to detect the presence of  
 CC A. baumannii and other Acinetobacter species in a sample, in screening  
 CC compounds for the ability to interfere with the A. baumannii life cycle  
 CC or to inhibit A. baumannii infection, and as biocontrol agents for  
 CC plants. The present sequence represents the amino acid sequence of an A.  
 CC baumannii protein.  
 XX SQ Sequence 729 AA;  
 Query Match 79.5%; Score 31; DB 6; Length 729;  
 Best Local Similarity 62.5%; Pred. No. 6.3e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 Db 279 GIGVSIRP 286  
 RESULT 9  
 ADD31284  
 ID ADD31284 standard; peptide; 8 AA.  
 XX AC ADD31284;  
 XX DT 15-JAN-2004 (first entry)  
 XX DE Angiogenesis inhibiting peptide #59.  
 XX KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT Modified-site 3  
 FT /label= OTHER  
 FT /note= "OTHER= D-form residue, alle (not defined)"  
 FT Modified-site 8  
 FT /note= "NHCH2CH3"  
 XX US2003109455-A1.  
 XX PN 12-JUN-2003.  
 XX PD 30-OCT-2002; 2002US-00283550.  
 XX PF 31-OCT-2001; 2001US-0335017P.  
 XX PR (HAVI/) HAVIV F.  
 XX PA (BRAD/) BRADLEY M F.  
 XX PI Haviv F, Bradley MF;  
 XX DR WPI; 2003-843101/78.  
 XX PT New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.  
 XX Claim 13; Page 25; 26pp; English.  
 PS CC The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX SQ Sequence 8 AA;  
 Query Match 76.9%; Score 30; DB 7; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 Db 1 GVXTSIRP 8  
 RESULT 10  
 ADE83635  
 ID ADE83635 standard; peptide; 8 AA.  
 XX AC ADE83635;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Antiangiogenic peptide.  
 XX KW antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiac; vulnerary;  
 KW antiinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Weber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiodiroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.  
 XX OS

OS Synthetic.  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "acetylated"  
 FT Modified-site 3 /note= "D-form residue alloisoleucyl"  
 FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"  
 FT  
 XX WO2003037268-A2.  
 XX  
 XX 08-MAY-2003.  
 XX  
 XX 30-OCT-2002; 2002WO-US034811.  
 XX  
 XX 31-OCT-2001; 2001US-00000681.  
 PR 04-OCT-2002; 2002US-00263812.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Bradley MF;  
 XX  
 XX WPI; 2003-617886/58.  
 DR  
 XX New hepta-, octa- and nona-peptide compounds used for treating e.g.  
 PT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic  
 PT retinopathy.  
 XX  
 XX Claim 13; Page 46; 5lpp; English.  
 XX  
 CC The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antiatherosclerotic,  
 CC cardiant, vulnary, antiinflammatory, antitumor and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumors. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumors. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.  
 XX  
 XX Sequence 8 AA;  
 XX  
 CC Query Match 76.9%; Score 30; DB 7; Length 8;  
 CC Best Local Similarity 75.0%; Pred. No. 1.4e-06;  
 CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 DB |||||  
 1 GVXTSIRP 8  
 XX  
 XX AAU15042;  
 XX  
 XX 04-DEC-2001 (first entry)  
 XX  
 DE Antiangiogenic peptide #69 useful for inhibiting angiogenesis.  
 XX  
 XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW

KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /label= Megly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT  
 FT Modified-site 4 /note= "D-form residue, alloisoleucine"  
 FT Modified-site 6 /note= "N-Methyl serine"  
 FT Modified-site 9 /note= "Modified by NH-ethyl"  
 FT  
 XX WO200138397-A1.  
 XX  
 XX 31-MAY-2001.  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 XX  
 XX 22-NOV-1999; 99US-00447099.  
 PR 31-OCT-2000; 2000US-00702649.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.  
 XX  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 XX diabetic retinopathy.  
 XX  
 XX Claim 39; Page 89; 95pp; English.  
 XX  
 CC The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #69  
 XX  
 XX Sequence 9 AA;  
 XX  
 CC Query Match 76.9%; Score 30; DB 4; Length 9;  
 CC Best Local Similarity 75.0%; Pred. No. 1.4e-06;  
 CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 DB |||||  
 2 GVXTSIRP 9  
 XX  
 XX RESULT 12  
 XX AAU14987  
 ID AAU14987 standard; peptide; 10 AA.

XX AAU14987;  
 XX AC  
 XX DT 04-DEC-2001 (first entry)  
 XX DE Antiangiogenic peptide #14 useful for inhibiting angiogenesis.  
 XX KW Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 XX KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 XX KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 XX KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /label= Megly  
 XX FT /note= "N-methyl-glycine (sarcosine), additionally  
 XX FT modified by N-terminal acetyl"  
 XX FT 4  
 XX FT Modified-site /note= "D-form residue, alioleucine"  
 XX FT 6  
 XX FT Modified-site /note= "N-Methyl serine"  
 XX FT 10  
 XX FT Modified-site /note= "D-form residue, C-terminal amide"  
 XX WO200138397-A1.  
 XX 31-MAY-2001.  
 XX 22-NOV-2000; 2000WO-US032105.  
 XX 22-NOV-1999; 99US-00447099.  
 XX 31-OCT-2000; 2000US-00702649.  
 XX (ABBO ) ABBOTT LAB.  
 XX Havig F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 XX diabetic retinopathy.  
 XX Claim 38; Page 88; 95pp; English.  
 XX The present invention relates to novel synthetic antiangiogenic peptides  
 XX (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 XX antiangiogenic peptides are useful for isolating a receptor from an  
 XX endothelial cell. The peptides of the invention are also useful for  
 XX treating cancer, arthritis, psoriasis and other skin diseases,  
 XX angiogenesis of the eye associated with infection or surgical  
 XX intervention and other ocular diseases, cat scratch disease, ulcers,  
 XX macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 XX haemangiomas and capillary action within atherosclerotic plaques,  
 XX autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 XX Osler-Webber syndrome, myocardiastasia, haemophilic joints, angiofibroma,  
 XX neovascularisation, telangiectasia, plaque  
 XX wound granulation, excessive or abnormal stimulation of endothelial  
 XX cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 XX scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 XX peptides are also useful as birth control agents by inhibiting ovulation  
 XX and to reduce bleeding by administration before surgery. The peptides of  
 XX the invention exhibit improved metabolic stability, improved oral  
 XX pharmacokinetics, increased water solubility, and improved oral  
 XX availability. The present sequence represents antiangiogenic peptide #14  
 XX Sequence 10 AA;  
 XX Query Match 76.9%; Score 30; DB 4; Length 10;  
 XX Best Local Similarity 75.0%; Pred. No. 16;  
 XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 |||||  
 Db 2 GVXTSIRP 9  
 RESULT 13  
 AAB35743  
 ID AAB35743 standard; protein; 12 AA.  
 XX AC AAB35743;  
 XX 20-FEB-2001 (first entry)  
 XX Glucodextranase peptide fragment GD LY-4.  
 XX Glucodextranase; isomalto-oligosaccharide; production.  
 XX Glucodextranase; isomalto-oligosaccharide; production.  
 XX Arthrobacter globiformis.  
 XX JP2000245475-A.  
 XX 12-SEP-2000.  
 XX 03-MAR-1999; 99JP-00056129.  
 XX 03-MAR-1999; 99JP-00056129.  
 XX (NISO ) NIPPON SHOKUHIN KAKO KK.  
 XX WPI; 2000-615145/59.  
 XX Glucodextranase gene and preparation of a polypeptide having  
 XX glucodextranase activity.  
 XX Example 1; Page 5; 15pp; Japanese.  
 XX This invention relates to a glucodextranase gene, the protein it encodes.  
 XX The invention includes a method for the preparation of a polypeptide  
 XX having glucodextranase activity in which a microbe transformed by a  
 XX vector containing the gene is cultured and a polypeptide having  
 XX glucodextranase activity is collected. The polypeptide can be used for  
 XX the preparation of an isomalto-oligosaccharide. The present sequence  
 XX represents a peptide fragment of the glucodextranase protein of the  
 XX invention  
 XX Sequence 12 AA;  
 XX Query Match 76.9%; Score 30; DB 3; Length 12;  
 XX Best Local Similarity 62.5%; Pred. No. 19;  
 XX Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 |||||  
 Db 2 GVGTSLNP 9  
 RESULT 14  
 AAU49108  
 ID AAU49108 standard; protein; 64 AA.  
 XX AC AAU49108;  
 XX 27-FEB-2002 (first entry)  
 XX Propionibacterium acnes immunogenic protein #10004.  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 XX dermatological; osteopathic; neuroprotectant.  
 XX Propionibacterium acnes.  
 XX OS

```

XX PN WO200181581-A2.
XX PD 01-NOV-2001.
XX PF
XX PF 20-APR-2001; 2001WO-US012865.
XX PR 21-APR-2000; 2000US-0199047P.
XX PR 02-JUN-2000; 2000US-0208841P.
XX PR 07-JUL-2000; 2000US-0216747P.
XX PA (CORI-) CORIXA CORP.
XX PI
XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX DR N-PSDB; AAS59544.
XX DR
XX DR
XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX PS
XX PS Example 1; SEQ ID NO 10303; 1069pp; English.
XX CC
XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX CC polypeptides. The proteins and their associated DNA sequences are used in
XX CC the treatment, prevention and diagnosis of medical conditions caused by
XX CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
XX CC P. acnes is also involved in infections of bone, joints and the central
XX CC nervous system, however it is particularly involved in the inflammatory
XX CC lesions associated with acne vulgaris. A method for detecting the
XX CC presence or absence of P. acnes in a patient comprises contacting a
XX CC sample with a binding agent that binds to the proteins of the invention
XX CC and determining the amount of bound protein in the sample. The
XX CC polypeptides may be used as antigens in the production of antibodies
XX CC specific for P. acnes proteins. These antibodies can be used to
XX CC downregulate expression and activity of P. acnes polypeptides and
XX CC therefore treat P. acnes infections. The antibodies may also be used as
XX CC diagnostic agents for determining P. acnes presence, for example, by
XX CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX CC this patent did not form part of the printed specification, but was
XX CC obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX SQ Sequence 64 AA;
XX
XX Query Match 76.9%; Score 30; DB 4; Length 64;
XX Best Local Similarity 62.5%; Pred. No. 97;
XX Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0
XX
XX QY 2 GVGTXIRP 9
XX DB |||: |||
XX 44 GVGSQLRP 51
XX
XX RESULT 15
XX ABM45627
XX ID ABM45627 standard; protein; 64 AA.
XX AC
XX AC ABM45627;
XX DT
XX DT 20-OCT-2003 (first entry)
XX DE
XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #10303.
XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX KW immunostimulant; immune response; vaccine.
XX OS
XX OS Propionibacterium acnes.
XX PN WO2003033515-A1.

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:30:38 ; Search time 12.7895 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-3

Perfect score: 39

Sequence: 1 XGVGTIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	89.7	271	4	US-09-252-991A-26235
2	32	82.1	307	4	US-09-252-991A-19676
3	31	79.5	95	3	US-09-024-023-4
4	31	79.5	95	4	US-09-531-111-4
5	31	79.5	229	3	US-09-024-023-2
6	31	79.5	229	4	US-09-531-111-2
7	31	79.5	397	4	US-09-252-991A-18356
8	31	79.5	727	4	US-09-540-236-3023
9	31	79.5	729	4	US-09-328-352-8036
10	30	76.9	123	4	US-09-732-210-76
11	30	76.9	241	3	US-08-834-776A-2
12	30	76.9	335	3	US-08-736-915-2
13	30	76.9	581	4	US-09-252-991A-20966
14	30	76.9	2318	3	US-09-091-219-24
15	30	76.9	2318	4	US-09-660-541-24
16	30	76.9	2680	4	US-09-489-039A-7973
17	29	74.4	71	4	US-09-543-681A-4174
18	29	74.4	226	1	US-08-650-129-4
19	29	74.4	226	3	US-08-984-417-4
20	29	74.4	227	1	US-08-569-806-8
21	29	74.4	255	1	US-08-650-129-5
22	29	74.4	255	3	US-08-984-417-5
23	29	74.4	360	4	US-09-786-240-2
24	29	74.4	361	3	US-09-032-372-12
25	29	74.4	361	4	US-09-786-240-32
26	29	74.4	373	4	US-09-786-240-33
27	29	74.4	504	4	US-09-554-726A-10

28	74.4	510	4	US-09-252-991A-17695	Sequence 17695, A
29	74.4	510	4	US-09-252-991A-29467	Sequence 29467, A
30	74.4	879	4	US-09-252-991A-28118	Sequence 28118, A
31	74.4	3567	2	US-07-642-734C-4	Sequence 4, Appli
32	74.4	3567	3	US-08-439-009A-4	Sequence 4, Appli
33	71.8	53	4	US-09-800-729-165	Sequence 165, App
34	71.8	160	4	US-09-252-991A-31843	Sequence 31843, A
35	71.8	163	4	US-09-562-737-76	Sequence 76, Appl
36	71.8	232	4	US-09-328-352-8189	Sequence 8189, Ap
37	71.8	239	4	US-09-328-352-7328	Sequence 7328, Ap
38	71.8	254	4	US-09-252-991A-23311	Sequence 23311, A
39	71.8	261	3	US-08-483-857-10	Sequence 10, Appl
40	71.8	265	4	US-09-543-681A-8110	Sequence 8110, A
41	71.8	309	4	US-09-252-991A-17975	Sequence 17975, A
42	71.8	338	4	US-09-199-637A-47	Sequence 47, Appl
43	71.8	355	1	US-08-196-218-34	Sequence 34, Appl
44	71.8	355	1	US-08-681-953-34	Sequence 34, Appl
45	71.8	362	4	US-09-134-000C-3578	Sequence 3578, Ap

## ALIGNMENTS

RESULT 1  
US-09-252-991A-26235  
; Sequence 26235, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26235  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26235

Query Match 89.7%; Score 35; DB 4; Length 271;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 85 GVGTSIRP 92

RESULT 2  
US-09-252-991A-19676  
; Sequence 19676, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19676  
; LENGTH: 307  
; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-252-391A-19676

Query Match 82.1%; Score 32; DB 4; Length 307;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 23 GVGTTARP 30

RESULT 3  
US-09-024-023-4  
Sequence 4, Application US/09024023  
Patent No. 6110899  
GENERAL INFORMATION:  
APPLICANT: Lonetto, Michael A.  
TITLE OF INVENTION: licc  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch St  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/024,023  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/039,210  
FILING DATE: 28-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Falk, Stephen T  
REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM50018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-024-023-4

Query Match 79.5%; Score 31; DB 3; Length 95;  
Best Local Similarity 62.5%; Pred. No. 25;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 9 GLGTLRLP 16

RESULT 4  
US-09-531-111-4  
Sequence 4, Application US/09531111  
Patent No. 6303571  
GENERAL INFORMATION:  
APPLICANT: Lonetto, Michael A.  
TITLE OF INVENTION: licc  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch St  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/531,111  
FILING DATE: 17-Mar-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/024,023  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Falk, Stephen T  
REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM50018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 95 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-531-111-4

Query Match 79.5%; Score 31; DB 4; Length 95;  
Best Local Similarity 62.5%; Pred. No. 25;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 9 GLGTLRLP 16

RESULT 5  
US-09-024-023-2  
Sequence 2, Application US/09024023  
Patent No. 6110899  
GENERAL INFORMATION:  
APPLICANT: Lonetto, Michael A.  
TITLE OF INVENTION: licc  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic tower, 1717 Arch St  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/024,023  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/039,210  
FILING DATE: 28-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Falk, Stephen T

REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM50018  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX:

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

US-09-024-023-2

Query Match 79.5%; Score 31; DB 3; Length 229;  
Best Local Similarity 62.5%; Pred. No. 58;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 9 GLGTRLRP 16

RESULT 6

US-09-531-111-2  
Sequence 2, Application US/09531111  
Patent No. 6303571  
GENERAL INFORMATION:  
APPLICANT: Lonetto, Michael A.  
TITLE OF INVENTION: licc  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert, Price & Rhoads  
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103-2793

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
FILING DATE: 17-Mar-2000  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/024,023  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:

NAME: Falk, Stephen T  
REGISTRATION NUMBER: 36,795  
REFERENCE/DOCKET NUMBER: GM50018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-994-2488  
TELEFAX: 215-994-2222  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-531-111-2

Query Match 79.5%; Score 31; DB 4; Length 229;  
Best Local Similarity 62.5%; Pred. No. 58;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|

Db 9 GLGTRLRP 16

RESULT 7

US-09-252-991A-18356  
Sequence 18356, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252.991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 18356  
LENGTH: 397  
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18356

Query Match 79.5%; Score 31; DB 4; Length 397;  
Best Local Similarity 75.0%; Pred. No. 99;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 136 GVGTVTRP 143

RESULT 8

US-09-540-236-3023  
Sequence 3023, Application US/09540236  
Patent No. 6673910  
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT  
FILE REFERENCE: 2709.2005-001  
CURRENT APPLICATION NUMBER: US/09/540.236  
CURRENT FILING DATE: 2000-04-04  
NUMBER OF SEQ ID NOS: 3840

SEQ ID NO 3023  
LENGTH: 727  
TYPE: PRT

ORGANISM: M.catarrhalis  
US-09-540-236-3023

Query Match 79.5%; Score 31; DB 4; Length 727;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 276 GIGVSIRP 283

RESULT 9

US-09-328-352-8036  
Sequence 8036, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: BAUMANNI FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-03FA  
CURRENT APPLICATION NUMBER: US/09/328.352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252

QY 2 GVGTXIRP 9  
|:|:|:|



; SEQ ID NO 8036  
; LENGTH: 729  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-8036

Query Match 79.5%; Score 31; DB 4; Length 729;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|:|  
Db 279 GIGVSIKP 286

RESULT 10  
US-09-732-210-76  
; Sequence 76, Application US/09732210  
; Patent No. 6573361  
; GENERAL INFORMATION:  
; APPLICANT: Bunkers, Jhong  
; APPLICANT: Liang, Jihong  
; APPLICANT: Mittanck, Cindy A.  
; APPLICANT: Seale, Jeffrey W.  
; APPLICANT: Wu, Yornie S.  
; TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use  
; FILE REFERENCE: 38-21(15036)B  
; CURRENT APPLICATION NUMBER: US/09732,210  
; CURRENT FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,513  
; PRIOR FILING DATE: 1999-12-07  
; PRIOR APPLICATION NUMBER: US 60/169,340  
; PRIOR FILING DATE: 1999-12-07  
; NUMBER OF SEQ ID NOS: 1753  
; SEQ ID NO 76  
; LENGTH: 123  
; TYPE: PRT  
; ORGANISM: Halobacterium halobium  
US-09-732-210-76

Query Match 76.9%; Score 30; DB 4; Length 123;  
Best Local Similarity 62.5%; Pred. No. 52;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|:|  
Db 76 GVGTAFKP 83

RESULT 11  
US-08-834-776A-2  
; Sequence 2, Application US/08834776A  
; Patent No. 6060241  
; GENERAL INFORMATION:  
; APPLICANT: Cortesey-Theulaz, Irene  
; TITLE OF INVENTION: Compositions and Methods Relating to  
; TITLE OF INVENTION: Drug Discovery and Detection and Treatment of  
; TITLE OF INVENTION: Gastrointestinal Diseases  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
; STREET: Five Palo Alto Square, 3000 El Camino Real  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306-2155  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/834,776A

; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ashton, Nina M.  
; REGISTRATION NUMBER: 37,273  
; REFERENCE/DOCKET NUMBER: GAST-001/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-843-5000  
; TELEFAX: 415-857-0663  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-834-776A-2

Query Match 76.9%; Score 30; DB 3; Length 241;  
Best Local Similarity 75.0%; Pred. No. 98;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|:|  
Db 130 GVGTLIAP 137

RESULT 12  
US-08-736-915-2  
; Sequence 2, Application US/08736915  
; Patent No. 6180385  
; GENERAL INFORMATION:  
; APPLICANT: Ahorn, Horst  
; APPLICANT: Blaas, Dieter  
; APPLICANT: Kirchweyer, Regina  
; APPLICANT: Lamphear, Barry J.  
; APPLICANT: Liebig, Hans-Dieter  
; APPLICANT: Rhoads, Robert E.  
; APPLICANT: Skern, Timothy  
; APPLICANT: Sommergruber, Wolfgang  
; APPLICANT: Waters, Debra  
; APPLICANT: Ziegler, Elisabeth  
; TITLE OF INVENTION: Picornavirus L Proteinase and Methods of  
; TITLE OF INVENTION: Making and Using Thereof  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX  
; STREET: 1100 New York Ave., NW  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/736,915  
; FILING DATE: 25-OCT-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/279,152  
; FILING DATE: 22-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 0652.1390000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-2600  
; TELEFAX: 202-371-2540  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 335 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; IMMEDIATE SOURCE: Fig. 1  
US-08-736-915-2

Query Match 76.9%; Score 30; DB 3; Length 335;  
Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 86 GIGTASRP 93

RESULT 13  
US-09-252-991A-20966  
; Sequence 20966, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20966  
; LENGTH: 581  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20966

Query Match 76.9%; Score 30; DB 4; Length 581;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 484 GPGTAVRP 491

RESULT 14  
US-09-091-219-24  
; Sequence 24, Application US/09091219  
; Patent No. 6171592  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/091,219  
; CURRENT FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER APPLICATION NUMBER: AU PN7201  
; EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-091-219-24

Query Match 76.9%; Score 30; DB 3; Length 2318;

Best Local Similarity 62.5%; Pred. No. 8.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 114 GIGTASRP 121

RESULT 15  
US-09-660-541-24  
; Sequence 24, Application US/09660541  
; Patent No. 6531136  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/660,541  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201  
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-660-541-24

Query Match 76.9%; Score 30; DB 4; Length 2318;  
Best Local Similarity 62.5%; Pred. No. 8.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
|:|:|:|  
Db 114 GIGTASRP 121

Search completed: April 1, 2004, 17:42:12  
Job time : 13.7895 secs



```

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 161814
; LENGTH: 198
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(198)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117135C.1.pep
; US-10-424-599-161814

Query Match          79.5%; Score 31; DB 12; Length 198;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
DB      58 GVGTAPEP 65

RESULT 3
US-10-425-114-43859
; Sequence 43859, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43859
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700657218_FLI.pep
; US-10-425-114-43859

Query Match          79.5%; Score 31; DB 12; Length 277;
Best Local Similarity 62.5%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
DB      71 GVGTAPEP 78

RESULT 4
US-10-282-122A-57397
; Sequence 57397, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA_034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 57397
; LENGTH: 281
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; US-10-282-122A-57397

Query Match          79.5%; Score 31; DB 12; Length 281;
Best Local Similarity 75.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
DB      59 GVGTYIEP 66

RESULT 5
US-10-369-493-9133
; Sequence 9133, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 9133
; LENGTH: 308
; TYPE: PRT
```

; ORGANISM: Chloroflexus aurantiacus  
US-10-369-493-9133

Query Match 79.5%; Score 31; DB 15; Length 308;  
Best Local Similarity 62.5%; Pred. No. 1.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 9 GLGTLRLP 16  
||| :|

RESULT 6  
US-10-369-493-23443  
; Sequence 23443, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 23443  
; LENGTH: 361  
; TYPE: PRT  
; ORGANISM: Deinococcus radiodurans  
US-10-369-493-23443

Query Match 79.5%; Score 31; DB 15; Length 361;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 18 GLGTLRLP 25  
||| :|

RESULT 7  
US-10-369-493-9512  
; Sequence 9512, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9512  
; LENGTH: 415  
; TYPE: PRT  
; ORGANISM: Xylella fastidiosa  
US-10-369-493-9512

Query Match 79.5%; Score 31; DB 15; Length 415;  
Best Local Similarity 62.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 83 GVGTLRLP 90  
||| :|

RESULT 8  
US-10-369-493-9289  
; Sequence 9289, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 9289  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Xylella fastidiosa  
US-10-369-493-9289

Query Match 79.5%; Score 31; DB 15; Length 416;  
Best Local Similarity 62.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 92 GVGTLRLP 99  
||| :|

RESULT 9  
US-10-369-493-17526  
; Sequence 17526, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 17526  
; LENGTH: 437  
; TYPE: PRT  
; ORGANISM: Xylella fastidiosa  
US-10-369-493-17526

Query Match 79.5%; Score 31; DB 15; Length 437;  
Best Local Similarity 62.5%; Pred. No. 2.7e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
DB 98 GVGTLRLP 105  
||| :|

RESULT 10  
US-10-156-761-11464

```
; Sequence 11464, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION: SATOSHI
; APPLICANT: OKURA, HARUO
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11464
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11464

Query Match          79.5%; Score 31; DB 14; Length 855;
Best Local Similarity 62.5%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
Db      703 GAGTLVRP 710

RESULT 11
US-10-120-801-81
; Sequence 81, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shalomit
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Komives, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 81
; LENGTH: 2043
; TYPE: PRT
; ORGANISM: Geodia cydonium
US-10-120-801-81

Query Match          79.5%; Score 31; DB 15; Length 2043;
Best Local Similarity 62.5%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
Db      71 GFGTAVRP 78

RESULT 12
US-10-424-599-260536
; Sequence 260536, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260536
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77287C.1.pep
US-10-424-599-260536

Query Match          76.9%; Score 30; DB 12; Length 200;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVGTXIRP 9
      |||||
Db      20 GRGTTVRP 27

RESULT 13
US-10-156-761-14597
; Sequence 14597, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OKURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14597
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14597
```

Query Match 76.9%; Score 30; DB 14; Length 253;  
Best Local Similarity 62.5%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVGTXIRP 9  
| | | | : |  
Db 70 GVGTTVAP 77

## RESULT 14

US-10-282-122A-69930  
; Sequence 69930, Application US/10282122A  
; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 69930

; LENGTH: 259

; TYPE: PRT

; ORGANISM: Pseudomonas syringae

US-10-282-122A-69930

Query Match 76.9%; Score 30; DB 12; Length 259;  
Best Local Similarity 62.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVGTXIRP 9  
| | | | : |  
Db 44 GVGTVNPF 51

## RESULT 15

US-10-369-493-10888

; Sequence 10888, Application US/10369493

; Publication No. US20030233675A1

## GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 10888

; LENGTH: 351

; TYPE: PRT

; ORGANISM: Ferroplasma acidamarum

US-10-369-493-10888

Query Match 76.9%; Score 30; DB 15; Length 351;

Best Local Similarity 62.5%; Pred. No. 3.5e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVGTXIRP 9

| | | | : |

Db 9 GAGTFLRP 16

Search completed: April 1, 2004, 17:45:38

Job time : 30.6316 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 9.78947 Seconds  
(without alignments)  
88.434 Million cell updates/sec

Title: US-09-833-196-3

Perfect score: 39

Sequence: 1 XGVGTXRIP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR: 78:\*

2: PIR: 1:\*

3: PIR: 3:\*

4: PIR: 4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	89.7	53	2 H84193	hypothetical prote
2	33	84.6	340	2 G70741	hypothetical prote
3	32	82.1	347	2 A95341	probable hydrolase
4	32	82.1	462	2 S74579	carboxyl-terminal
5	31	79.5	229	2 B95147	licC protein [impo
6	31	79.5	229	2 H98014	hypothetical prote
7	31	79.5	353	2 T46521	Probable dTDP-1-gl
8	31	79.5	361	2 A75596	Probable glucose-1
9	31	79.5	366	2 F83281	cobalamin biosynth
10	31	79.5	377	2 C82876	oligopeptide trans
11	31	79.5	437	2 G82777	glutamate symport
12	31	79.5	747	2 G87402	hypothetical prote
13	31	79.5	2043	2 T18524	scavenger receptor
14	30	76.9	123	2 T43824	hypothetical prote
15	30	76.9	123	2 D84322	hypothetical prote
16	30	76.9	217	2 G84045	transcription regu
17	30	76.9	242	2 F87687	nucleosidyltransfe
18	30	76.9	248	2 A97362	salF-like protein
19	30	76.9	248	2 AB2580	conserved hypotet
20	30	76.9	273	2 T48521	hypothetical prote
21	30	76.9	275	2 AD2333	hypothetical prote
22	30	76.9	281	2 B90167	conserved hypotet
23	30	76.9	315	2 C87293	hypothetical prote
24	30	76.9	352	2 B69290	glucose-1-phosphat
25	30	76.9	352	2 A26994	strD protein - Str
26	30	76.9	355	2 T30872	dNDP-glucose synth
27	30	76.9	385	2 G69101	mannose-1-phosphat
28	30	76.9	436	2 AI1015	probable exported
29	30	76.9	463	2 F70627	Probable narX prot

30	30	76.9	470	2 AB1804	transmembrane effl
31	30	76.9	470	2 AD1430	transmembrane effl
32	30	76.9	617	2 S38923	hypothetical prote
33	30	76.9	629	2 AB1860	acetolactate synth
34	30	76.9	843	2 S76785	hypothetical prote
35	30	76.9	1011	1 GNNYC1	genome polyprotein
36	30	76.9	2332	1 GNNYF	genome polyprotein
37	30	76.9	2332	1 GNNY4F	genome polyprotein
38	30	76.9	2333	1 GNNY2P	genome polyprotein
39	30	76.9	2336	2 S37077	mannose-1-phosphat
40	29	74.4	23	2 A47415	hypothetical prote
41	29	74.4	203	2 F87540	brachyurin (SC 3.4
42	29	74.4	226	1 KCUF	probable GntR-type
43	29	74.4	243	2 F95288	acetylglutamate ki
44	29	74.4	290	2 G69409	sporulation transc
45	29	74.4	295	2 T35193	

#### ALIGNMENTS

##### RESULT 1

H84193

hypothetical protein Vng0346h [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: H84193

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: H84193

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-53 <STO>

A:Cross-references: GB:AE004437; NID:G10579970; PIDN:ARG18916.1; GSPDB:GN00138

C:Genetics:

A:Gene: VNG0346H

Query Match	89.7%;	Score 35;	DB 2;	Length 53;
Best Local Similarity	75.0%;	Pred. No. 0.97;		
Matches	6;	Conservative	1;	Mismatches
				Indels
				Gaps
				0;

QY 2 GVGTXIRP 9

Db 34 GVGTVGRP 41

##### RESULT 2

G70741

hypothetical protein Rv1360 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: G70741

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: G70741

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-340 <COL>

A:Cross-references: GB:Z75555; GB:AL123456; NID:G3261608; PIDN:CAA99965.1; PID:e250558;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: Rv1360



Query Match 84.6%; Score 33; DB 2; Length 340;  
Best Local Similarity 75.0%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|:|:|:|  
DB 93 GLGTQIRP 100

RESULT 3  
A:Species: probable hydrolase protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: A95341  
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Kemp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: A95341  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-347 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AAK65291.1; PID:gl4523746; GSPDB:GNC0165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhrer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Sma1166  
A:Genome: plasmid

Query Match 82.1%; Score 32; DB 2; Length 347;  
Best Local Similarity 62.5%; Pred. No. 29;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|:|:|:|  
DB 328 GAGTSVRP 335

RESULT 4  
S:4579  
carboxyl-terminal proteinase ctpB - Synecocystis sp. (strain PCC 6803)  
N:Alternate names: hypothetical protein slr0257  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997  
C:Accession: S74579  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis  
s.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74579  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-462 <KAN>  
A:Cross-references: EMBL:D90500; GB:AB001339; NID:gl651768; PIDN:BAAL6731.1; PID:gl65180  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: ctpB  
C:Superfamily: carboxyl-terminal processing proteinase  
Query Match 82.1%; Score 32; DB 2; Length 462;

Best Local Similarity 62.5%; Pred. No. 38;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|:|:|:|  
DB 20 GLGTALRP 27

RESULT 5  
B95147  
licC protein [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: B95147  
R:Tetellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei  
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfale,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison  
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: B95147  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK75371.1; PID:gl4972750; GSPDB:GN00164; TIGR:SP  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP1267

Query Match 79.5%; Score 31; DB 2; Length 229;  
Best Local Similarity 62.5%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|:|:|:|  
DB 9 GLGTALRP 16

RESULT 6  
H98014  
hypothetical protein licC [imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: H98014  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.;  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.;  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5703-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: H98014  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-229 <KUR>  
A:Cross-references: GB:AE007317; PIDN:AAK99948.1; PID:gl5458774; GSPDB:GN00174  
C:Genetics:  
A:Gene: licC

Query Match 79.5%; Score 31; DB 2; Length 229;  
Best Local Similarity 62.5%; Pred. No. 32;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
|:|:|:|  
DB 9 GLGTALRP 16

RESULT 7  
T46521  
probable drdp-1-glucose synthase [imported] - Streptomyces violaceoruber  
C:Species: Streptomyces violaceoruber

C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 11-May-2000  
 C;Accession: T46521  
 R;Ichinose, K.; Bedford, D.J.; Tornus, D.; Bechthold, A.; Bibb, M.J.; Revill, W.P.; Floss  
 Chen. Biol. 5, 647-659, 1998  
 A;Title: The granaticin biosynthetic gene cluster of Streptomyces violaceoruber Tu22: se  
 A;Reference number: Z23045; MUID:99051446; PMID:9831526  
 A;Accession: T46521  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-353 <ICH>  
 A;Cross-references: EMBL:AJ011500; PIDN:CAR09637.1  
 A;Experimental source: strain Tu22  
 C;Genetics:  
 A;Note: gra-orf16  
 C;Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 79.5%; Score 31; DB 2; Length 353;  
 Best Local Similarity 62.5%; Pred. No. 49;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 : : : : :  
 : : : : :  
 Db 9 GVGTXIRP 16

RESULT 8  
 A75596  
 C;Species: Deinococcus radiodurans - Deinococcus radiodurans (strain R1)  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C;Accession: A75596  
 R;White, O.; Bisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, L.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A;Reference number: A75250; MUID:20036896; PMID:10567266  
 A;Accession: A75596  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-361 <WHI>  
 A;Cross-references: GB:AB001862; GB:AB001825; NID:G6460468; PIDN:AAF12277.1; PID:G646057  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Gene: DRA0031  
 A;Map position: 2  
 C;Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 79.5%; Score 31; DB 2; Length 361;  
 Best Local Similarity 62.5%; Pred. No. 50;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 : : : : :  
 : : : : :  
 Db 18 GLGTRLRP 25

RESULT 9  
 F83281  
 Cobalamin biosynthetic protein CbID PA2908 [imported] - Pseudomonas aeruginosa (strain H  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: F83281  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-366 <STO>

A;Cross-references: GB:AB004717; GB:AB004091; NID:G9948999; PIDN:AA062296.1; GSPDB:GN00  
 A;Experimental source: strain PA01  
 C;Genetics:  
 A;Gene: cbid; PA2908  
 C;Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 79.5%; Score 31; DB 2; Length 366;  
 Best Local Similarity 75.0%; Pred. No. 50;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 : : : : :  
 : : : : :  
 Db 105 GVGTXIRP 112

RESULT 10  
 C82876  
 C;Species: Ureaplasma urealyticum  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C;Accession: C82876  
 R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.  
 Submitted to GenBank, February 2000  
 A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
 A;Reference number: A82870  
 A;Accession: C82876  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-377 <GLA>  
 A;Cross-references: GB:AB002154; GB:AF22894; NID:G6899557; PIDN:AAF30975.1; GSPDB:GN00  
 A;Experimental source: serovar 3; biovar 1  
 C;Genetics:  
 A;Gene: oppB; UU562  
 A;Genetic code: SGC3

Query Match 79.5%; Score 31; DB 2; Length 377;  
 Best Local Similarity 75.0%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 : : : : :  
 : : : : :  
 Db 146 GVGTXIRP 153

RESULT 11  
 G82777  
 Glutamate symport protein XF0656 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C;Accession: G82777  
 R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ  
 Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A;Reference number: A82515; MUID:20365717; PMID:10910347  
 A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: G82777  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-437 <SIM>  
 A;Cross-references: GB:AB003910; GB:AB003849; NID:G9105532; PIDN:AAF83466.1; GSPDB:GN00  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.R.; Kuramae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Sa, R.G.; Santelli, R.V.; Sawasa  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; da Silva, F.R.; da Silva, W.A.; da Silve  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, W.A.; da Silva, W.A.; da Silve  
 M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0656  
 C;Superfamily: Bacillus subtilis sodium-glutamate symporter homolog yhcL

Query Match 79.5%; Score 31; DB 2; Length 437;  
 Best Local Similarity 62.5%; Pred. No. 60;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 ||||| :  
 Db 98 GVGTLIRP 105

## RESULT 12

G87402  
 hypothetical protein CC1238 [imported] - Caulobacter crescentus  
 C;Species: Caulobacter crescentus  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C;Accession: G87402  
 R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A;Title: Complete Genome Sequence of Caulobacter crescentus.  
 A;Reference number: A87249; MUID:21173698; PMID:11259647  
 A;Accession: G87402  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-747 <STO>  
 A;Cross-references: GB:AB005673; NID:g13422565; PIDN:AAK23219.1; GSPDB:GN00148  
 C;Genetics:  
 A;Gene: CC1238

Query Match 79.5%; Score 31; DB 2; Length 747;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 ||||| :  
 Db 204 GVGKIRP 211

## RESULT 13

T18524  
 scavenger receptor cysteine-rich protein homolog srcrm2 - Geodia cydonium  
 C;Species: Geodia cydonium  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-May-2000  
 R;Blumbach, B.; Pancer, Z.; Diehl-Seifert, B.; Steffen, R.; Muenkner, J.; Mueller, I.; M  
 J. Cell Sci. 111, 2635-2644, 1998  
 A;Title: The putative sponge aggregation receptor: Isolation and characterisation of a m  
 A;Reference number: Z18947; MUID:98369060; PMID:9701562  
 A;Accession: T18524  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-2043 <BLU>  
 A;Cross-references: EMBL:Y14953; NID:e1364818; PID:e1364819; PIDN:CAA75175.1  
 C;Genetics:  
 A;Gene: SRCRM2

Query Match 79.5%; Score 31; DB 2; Length 2043;  
 Best Local Similarity 62.5%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 ||||| :  
 Db 71 GVGTAIRP 78

## RESULT 14

T43824

hypothetical protein [imported] - Halobacterium salinarum  
 C;Species: Halobacterium salinarum  
 C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 17-Nov-2000  
 C;Accession: T43824; S11600  
 R;Itch, T.  
 submitted to the EMBL Data Library, September 1997  
 A;Reference number: Z22697  
 A;Accession: T43824  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-123 <ITO>  
 A;Cross-references: EMBL:AB006961; PIDN:BA22278.1  
 A;Note: the source is designated as Halobacterium halobium  
 R;Spirdonova, V.A.; Akhmanova, A.S.; Kagramanova, V.K.; Koepke, A.K.E.; Mankin, A.S.  
 Can. J. Microbiol. 35, 153-159, 1989  
 A;Title: Ribosomal protein gene cluster of Halobacterium halobium: nucleotide sequence  
 A;Reference number: S11597; MUID:89248673; PMID:2470481  
 A;Accession: S11600  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 'M', 9-84 <SPI>  
 A;Note: the source is designated as Halobacterium halobium  
 C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0464

Query Match 76.9%; Score 30; DB 2; Length 123;  
 Best Local Similarity 62.5%; Pred. No. 28;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 ||||| :  
 Db 76 GVGTAIRP 83

## RESULT 15

D84322  
 hypothetical protein Vngi699c [imported] - Halobacterium sp. NRC-1  
 C;Species: Halobacterium sp. NRC-1  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 24-Aug-2001  
 C;Accession: D84322  
 R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
 ; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A;Authors: Hou, S.; Daniela, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; I  
 A;Title: Genome sequence of Halobacterium species NRC-1.  
 A;Reference number: A84160; MUID:20504483; PMID:11016950  
 A;Accession: D84322  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-123 <STO>  
 A;Cross-references: GB:AE004437; NID:gi0581167; PIDN:AAG19944.1; GSPDB:GN00138  
 C;Genetics:  
 A;Gene: VNGI699C  
 C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0464

Query Match 76.9%; Score 30; DB 2; Length 123;  
 Best Local Similarity 62.5%; Pred. No. 28;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
 ||||| :  
 Db 76 GVGTAIRP 83

Search completed: April 1, 2004, 17:40:40  
 Job time : 11.7895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: April 1, 2004, 17:29:18 ; Search time 5 52632 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-3  
Perfect score: 39  
Sequence: 1 XGVGTIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	84.6	340	1 YD60 MYCTU	Q11030 mycobacteri
2	32	82.1	923	1 AGLU_TETPY	O00906 tetrahymena
3	30	76.9	123	1 YG99 HALN1	O24785 halobacteri
4	30	76.9	355	1 STRD_STRGR	P08075 streptomyc
5	30	76.9	1011	1 POLG_FMDVT	P15072 foot-and-mo
6	30	76.9	1157	1 DD37 HUMAN	Q81Y37 homo sapien
7	30	76.9	2332	1 POLG_FMDVA	P03308 f genome po
8	30	76.9	2332	1 POLG_FMDV1	P03306 f genome po
9	30	76.9	2336	1 POLG_FMDV2	P49303 f genome po
10	30	76.9	2336	1 POLG_FMDV3	P00771 uca pugilac
11	29	74.4	226	1 CGGS_UCAPU	O52691 streptomyc
12	29	74.4	227	1 T2S1_STRCS	Q7nt63 chromobacte
13	29	74.4	289	1 ARGB_CHRFU	O28998 archaeglob
14	29	74.4	290	1 ARGB_ARCFU	P41940 saccharomyc
15	29	74.4	361	1 MPGL_YEAST	O74484 schizosacch
16	29	74.4	363	1 MPGL_SCHPO	Q9K993 bacillus ha
17	29	74.4	440	1 XFLA_BACHD	O09130 schizosacch
18	29	74.4	445	1 IF2G_SCHPO	P41091 homo sapien
19	29	74.4	471	1 IF2G_HUMAN	Q9z0n1 mus musculu
20	29	74.4	471	1 IF2G_MOUSE	Q9z0n2 mus musculu
21	29	74.4	471	1 IF2H_MOUSE	O24208 drosophila
22	29	74.4	475	1 IF2G_DROME	O03132 saccharopol
23	29	74.4	3567	1 ERY2_SACER	P08632 rhizobium l
24	28	71.8	190	1 NODL_RHILV	P75939 escherichia
25	28	71.8	260	1 FLGG_ECOLI	P16439 salmonella
26	28	71.8	260	1 FLGG_SALTY	P59295 bifidobacte
27	28	71.8	305	1 ARGB_BIFLO	Q74529 synecocyst
28	28	71.8	317	1 MURB_SYNY3	Q7mbc4 gloeobact
29	28	71.8	357	1 MURG_GLOVI	O58657 methanococc
30	28	71.8	437	1 IF2G_METJA	O887q2 pseudomon
31	28	71.8	493	1 ALGE_PSEPM	P32481 saccharomyc
32	28	71.8	527	1 IF2G_YEAST	Q99383 saccharomyc
33	28	71.8	534	1 NAB4_YEAST	

Q9pkv2 chlamydia m  
Q9rqi5 neisseria d  
Q9rdl3 corynebacte  
Q99557 ceratodon p  
P14712 arabidopsis  
P33530 nicotiana t  
P36773 lathyrus sa  
P15001 pisum sativ  
O49934 populus tre  
P95629 rhizobium m  
P25848 ceratodon p  
Q9y4c0 homo sapien

ALIGNMENTS

RESULT 1

YD60 MYCTU  
ID YD60 MYCTU STANDARD; PRT; 340 AA.  
AC Q11030;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical protein Rv1360/Mti1405/Mbi1395 precursor.  
GN Rv1360 OR Mti1405 OR Mty02B10.24 OR Mbi1395.  
OS Mycobacterium tuberculosis, and  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773, 1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
Davies R., Devlin K., Feldwell T., Gentles S., Hamlin N., Holroyd S.,  
Hornsbay T., Jagels K., Krogi A., McLean J., Moule S., Murphy L.,  
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence."  
RL Nature 393:537-544(1998).  
[2]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RA Fleischnann R.D., Allard D., Eisen J.A., Carpenter L., White O.,  
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains."  
RL J. Bacteriol. 184:5479-5490(2002).  
[3]  
RP SEQUENCE FROM N.A.  
RC SPECIES=M.bovis; STRAIN=AF2122/97;  
RX MEDLINE=22709107; PubMed=12788972;  
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,  
Pryor M., Dutthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;  
RT "The complete genome sequence of Mycobacterium bovis."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:17877-7882(2003).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; 275555; CA99965.1; -  
DR EMBL; AB007012; AAK45688.1; -  
DR EMBL; BX248338; CAD94256.1; -  
DR PIR; G70741; G70741.  
DR TIGR; MT1405; -

DR Tuberculin; Rv1360; -  
DR InterPro; IPR002103; Bac luciferase.  
KW Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 20 POTENTIAL  
FT CHAIN 21 340  
FT Rv1360/MT1405/MB1395.  
SQ SEQUENCE 340 AA; 37252 MW; F4D3BCA7C5AC0767 CRC64;

Query Match 84.6%; Score 33; DB 1; Length 340;  
Best Local Similarity 75.0%; Pred. No. 8.9; Indels 0; Gaps 0;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Dy 93 GLGTQIRP 100  
ID AGLU TETPY STANDARD; PRT; 923 AA.

RESULT 2  
AGLU TETPY  
AC O00906;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Lysosomal acid alpha-glucosidase precursor (EC 3.2.1.20) (Acid  
DE maltase).  
OS Tetrahymena pyriformis.  
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
OC Tetrahymenina; Tetrahymena.  
OX NCBI\_TaxID=5908;  
RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=W; PubMed=8768433;  
RX MEDLINE=96313350; PubMed=8768433;  
RA Alan S., Nakashima S., Devashiki Y., Banno Y., Hara A., Nozawa Y.;  
RT "Molecular cloning of a gene encoding acid alpha-glucosidase from  
RT Tetrahymena pyriformis."  
RL J. Eukaryot. Microbiol. 43:295-303(1996).  
CC -!- FUNCTION: ESSENTIAL FOR THE DEGRADATION OF GLYCOGEN TO GLUCOSE IN  
CC LYSOSOMES. HAS BOTH ALPHA-1,4 AND ALPHA-1,6-GLUCOSIDASE ACTIVITY.  
CC OPTIMUM ACTIVITY AT PH 4.0.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-  
CC linked D-glucose residues with release of D-glucose.  
CC -!- SUBCELLULAR LOCATION: LYSOSOMAL AND SECRETED.  
CC -!- SIMILARITY: Belongs to family 31 of glycosyl hydrolases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D83384; BAA20462.1; -  
DR InterPro; IPR000322; Glyco\_hydro\_31.  
DR Pfam; PF01055; Glyco\_hydro\_31; 1.  
DR PROSITE; PS00129; GLYCOSYL\_HYDROL\_F31\_1; 1.  
DR PROSITE; PS00707; GLYCOSYL\_HYDROL\_F31\_2; 1.  
KW Hydrolase; Glycosidase; Glycoprotein; Lysosome; Signal.  
FT SIGNAL 1 17 POTENTIAL  
FT PROPEP 18 36  
FT CHAIN 37 923  
FT ACT\_SITE 455  
FT LYSOSOMAL ACID ALPHA-GLUCOSIDASE.  
FT BY SIMILARITY.

FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 848 848 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 912 912 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 923 AA; 104116 MW; 6294809F43EA54C3 CRC64;

Query Match 82.1%; Score 32; DB 1; Length 923;  
Best Local Similarity 75.0%; Pred. No. 35; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Dy 689 GVGTVFRP 696  
ID YG99 HALN1 STANDARD; PRT; 123 AA.

RESULT 3  
YG99 HALN1  
ID YG99 HALN1 STANDARD; PRT; 123 AA.  
AC O24785; Q9HPC6;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein Vng1699C.  
GN VNG1699C.  
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and  
OS Halobacterium halobium.  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091, 2242;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=NRC-1;  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,  
RA Leitthaus B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
RN [2] SEQUENCE FROM N.A.  
RP SPECIES=H.halobium;  
RX MEDLINE=97031049; PubMed=8876975;  
RA Miyokawa T., Urayama T., Shimooka K., Itoh T.;  
RT "Organization and nucleotide sequences of ten ribosomal protein genes  
RT from the region equivalent to the S10 operon in the archaeobacterium,  
RT Halobacterium halobium."; 9:1209-1220(1996).  
RL Biochem. Mol. Biol. Int. 39:1209-1220(1996).  
CC -!- SIMILARITY: Belongs to the UPF0086 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE005077; AAG19944.1; -  
DR EMBL; AB006961; BAA22278.1; -  
DR PIR; D84322; D84322.  
DR PIR; T43824; T43824.  
DR Pfam; PF01868; UPF0086; 1.  
DR

DR ProDom: PD009163; UPF0086; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 123 AA; 13135 MW; E6941EE7E22237EB CRC64;  
  
Query Match 76.9%; Score 30; DB 1; Length 123;  
Best Local Similarity 62.5%; Pred. No. 17;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVGTXIRP 9  
| | | | |  
Db 76 GVGTAFAK 83  
  
RESULT 4  
STRD\_STRGR STANDARD; PRT; 355 AA.  
ID STED\_STRGR STANDARD; PRT; 355 AA.  
AC P08075; 13135 MW; E6941EE7E22237EB CRC64;  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glucose-1-phosphate thymidyltransferase (EC 2.7.7.24) (dUDP-glucose  
DE synthase) (dUDP-glucose pyrophosphorylase) (sugar-nucleotidylat  
DE enzyme).  
GN STED.  
OS Streptomyces griseus.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1911;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N2-3-11;  
RX MEDLINE=88040426; PubMed=3118332;  
RA Distler J., Ebert A., Mansouri K., Pissowatzki K., Stockmann M.,  
RA Papezberg W.;  
RT "Gene cluster for streptomycin biosynthesis in Streptomyces griseus:  
RT nucleotide sequence of three genes and analysis of transcriptional  
RT activity";  
RL Nucleic Acids Res. 15:8041-8056(1987).  
CC -!- FUNCTION: Involved in the biosynthesis of the streptose moiety of  
CC streptomycin.  
CC -!- CATALYTIC ACTIVITY: dTTP + alpha-D-glucose 1-phosphate =  
CC diphosphate + dUDP-glucose.  
CC -!- PATHWAY: Streptomycin biosynthesis.  
CC -!- SIMILARITY: Belongs to the glucose-1-phosphate  
CC thymidyltransferase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Y00459; CAA68514.1; -.  
DR PIR; A26984; A26984.  
DR InterPro; IPR005908; GIP thy trans s.  
DR InterPro; IPR005835; NTP transferase.  
DR Pfam; PF00483; NTP transferase; 1.  
DR TIGRfams; TIGR01208; tmlA\_long; 1.  
KW Streptomycin biosynthesis; Transferase; Kinase;  
KW Nucleotidyltransferase.  
SQ SEQUENCE 355 AA; 38098 MW; ED136AEA854EB7DA CRC64;  
  
Query Match 76.9%; Score 30; DB 1; Length 355;  
Best Local Similarity 62.5%; Pred. No. 42;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVGTXIRP 9  
| | | | |  
Db 9 GTGTRLRP 16

RESULT 5  
POLG\_FMDVT STANDARD; PRT; 1011 AA.  
ID POLG\_FMDVT STANDARD; PRT; 1011 AA.  
AC P15072; Q84755; Q84756; Q84757; Q84758;  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;  
DE Core protein P12] (Fragment).  
OS Foot-and-mouth disease virus (strain C1) (Aphthovirus C) (FMDV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=12121;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=84069809; PubMed=6316275;  
RA Beck E., Forbs S., Strebel K., Cattaneo R., Feil G.;  
RT "Structure of the FMDV translation initiation site and of the  
RT structural proteins";  
RL Nucleic Acids Res. 11:7873-7885(1983).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
CC each of which is composed of one copy each of proteins VP1, VP2,  
CC VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
CC VIRUS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X00130; CAA24960.2; -.  
DR PIR; A20288; GNNYCL.  
DR HSP; Q88571; LTME.  
DR InterPro; IPR004080; FMDVP1coat.  
DR InterPro; IPR008739; Peptidase\_C28.  
DR InterPro; IPR001676; Rhv.  
DR InterPro; IPR008975; Viral cap coat.  
DR Pfam; PF05408; Peptidase\_C28; 1.  
DR Pfam; PF00073; rhv; 3.  
DR PRINTS; PR01542; FMDVP1COAT.  
DR PolyProtein; Coat protein; Core protein; Nonstructural protein;  
KW Myristate; Lipoprotein.  
FT CHAIN 1 216 NONSTRUCTURAL PROTEIN P20A.  
FT CHAIN 217 286 COAT PROTEIN VP4.  
FT CHAIN 287 504 COAT PROTEIN VP2.  
FT CHAIN 505 723 COAT PROTEIN VP3.  
FT CHAIN 724 932 COAT PROTEIN VP1.  
FT CHAIN 933 >1011 CORE PROTEIN P12.  
FT LIPID 217 217 N-myristoyl glycine (in host)  
FT (By similarity).  
FT NON\_TER 1011 1011  
SQ SEQUENCE 1011 AA; 111535 MW; 8B8845600A560601 CRC64;  
  
Query Match 76.9%; Score 30; DB 1; Length 1011;  
Best Local Similarity 62.5%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVGTXIRP 9  
| | | | |  
Db 114 GIGTASRP 121  
  
RESULT 6  
DD37\_HUMAN STANDARD; PRT; 1157 AA.  
ID DD37\_HUMAN STANDARD; PRT; 1157 AA.

AC Q81Y37; Q9BU17; Q9P211;  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Probable ATP-dependent helicase DHX37 (DEAH-box protein 37).  
 GN DHX37 OR DDX37 OR KIAA1517.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Duodenum;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.L., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Wax S.J., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fanej J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Truchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE OF 178-1157 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20277482; PubMed=10819331;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.,  
 RT "Prediction of the coding sequences of unidentified human genes. XVII.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:143-150 (2000).  
 CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAH  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; BC002575; AA02575.1; ALT\_INIT.  
 DR EMBL; BC037964; AA037964.1; -  
 DR EMBL; AB040950; AA040950.1; -  
 DR Genbank; HGNC:17210; DHX37.  
 DR InterPro; IPR001410; DHX37.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR InterPro; IPR007502; Helicase\_dom.  
 DR Pfam; PF04408; HA2; 1.  
 DR Pfam; PF00271; Helicase\_C; 1.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC\_C; 1.  
 DR PROSITE; PS00690; DEAH\_ATP\_HELICASE; FALSE\_NEG.  
 DR KX Hydrolase; Helicase; ATP-binding.  
 FT NP BIND 275 282 ATP (POTENTIAL).  
 FT SITE 372 375 DEAH BOX.  
 FT CONFLICT 869 869 S -> G (IN REF. 1; AA02575).  
 FT CONFLICT 898 898 A -> S (IN REF. 2).  
 FT CONFLICT 1081 1081 R -> Q (IN REF. 1; AA02575).  
 FT CONFLICT 1130 1157 YLLAEYCEWLPQAMPEPDIKAWPTTTH -> CEFDQGGV

FT GVDKMGSLRQGLCALCTVSPGLAEGSGTAAQGLFAT (IN  
 FT REF. 2).  
 SQ SEQUENCE 1157 AA; 129544 MW; 49332175221B30C5 CRC64;  
 Query Match 76.9%; Score 30; DB 1; Length 1157;  
 Best Local Similarity 62.5%; Fred. No. 1.2e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVGTXIRP 9  
 DB 191 GVGTTVAP 198  
 RESULT 7  
 POLG\_FMDVA STANDARD; PRT; 2332 AA.  
 ID POLG\_FMDVA STANDARD; PRT; 2332 AA.  
 AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;  
 AC Q65044; Q65045; Q65046; Q65047;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked  
 DE proteins VP1 TO VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OX NCBI\_TaxID=12114;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85211015; PubMed=2987518;  
 RA Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,  
 RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.,  
 RT "Nucleotide and amino acid sequence coding for polypeptides of  
 RT foot-and-mouth disease virus type A12.";  
 RL J. Virol. 54:651-660 (1985).  
 RN [2]  
 RP SEQUENCE OF 1863-2332 FROM N.A.  
 RX MEDLINE=83225613; PubMed=6305004;  
 RA Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,  
 RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.,  
 RT "Identification of amino acid and nucleotide sequence of the  
 RT foot-and-mouth disease virus RNA polymerase.";  
 RL Virology 126:614-623 (1983).  
 RN [3]  
 RP SEQUENCE OF 715-955 FROM N.A.  
 RX MEDLINE=82061853; PubMed=6272395;  
 RA Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,  
 RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,  
 RA Bachrach H.L.,  
 RT "Cloned viral protein vaccine for foot-and-mouth disease: responses  
 RT in cattle and swine.";  
 RL Science 214:1125-1129 (1981).  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; M10975; AAA42593.1; -;  
 DR EMBL; J02187; AAA42670.1; -;  
 DR MEROPS; C03.008; -;  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR004080; FMDVpiccoat.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001676; Rnv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_P8.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF05408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rnv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR PRINTS; PR01542; FMDVP1COAT.  
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; Lipoprotein.  
 FT CHAIN 1 200 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 201 285 COAT PROTEIN VP4.  
 FT CHAIN 286 503 COAT PROTEIN VP2.  
 FT CHAIN 504 723 COAT PROTEIN VP3.  
 FT CHAIN 724 937 COAT PROTEIN VP1.  
 FT CHAIN 938 953 CORE PROTEIN X.  
 FT CHAIN 954 1107 CORE PROTEIN P14.  
 FT CHAIN 1108 1425 CORE PROTEIN P19.  
 FT CHAIN 1426 1578 CORE PROTEIN P19.  
 FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VEG1.  
 FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VEG2.  
 FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VEG3.  
 FT CHAIN 1650 1862 PROTEASE.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT LIPID 201 201 N-myristoyl Glycine (in host).  
 SQ SEQUENCE 2332 AA; 259408 MW; EE77DA739CBEDC6A CRC64;  
  
 Query Match 76.9%; Score 30; DB 1; Length 2332;  
 Best Local Similarity 62.5%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 GVGXIRP 9  
 DB 114 GIGTASRP 121  
  
 RESULT 8  
 ID POLG\_FMDVO STANDARD; PRT; 2332 AA.  
 AC P03305;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;  
 DE Core protein P12; Core protein P34; Core protein P14; Genome-linked  
 DE protein VEG; Protease (EC 3.4.22.-); RNA-directed RNA polymerase  
 DE (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain O1) (Aphthovirus O) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OX NCBI\_TaxID=73482;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O1K;  
 RA MEDLINE=84297249; PubMed=6089122;  
 RX Forbes S., Strebel K., Beck E., Schaller H.;  
 RT "Nucleotide sequence and genome organization of foot-and-mouth  
 RT disease virus.";  
 RL Nucleic Acids Res. 12:5587-6601(1984).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=O1BFS;  
 RA MEDLINE=83143292; PubMed=6298715;  
 RX Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;  
 RT "Comparison of the amino acid sequence of the major immunogen from  
 RT three serotypes of foot and mouth disease virus.";  
 RL Nucleic Acids Res. 10:8285-8295(1982).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RA MEDLINE=89143740; PubMed=2537470;  
 RX Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;  
 RT "The three-dimensional structure of foot-and-mouth disease virus at  
 RT 2.9-A resolution.";  
 RL Nature 337:709-716(1989).  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.  
 CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
 CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
 CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
 CC VIRUS.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X00871; CA25416.1; -;  
 DR EMBL; J02185; AAA42635.1; -;  
 DR PDB; 1QMY; 18-SEP-01.  
 DR PDB; 1QOL; 10-NOV-01.  
 DR MEROPS; C03.008; -;  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR004080; FMDVpiccoat.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001676; Rnv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_P8.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF05408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rnv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR PRINTS; PR01542; FMDVP1COAT.  
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; 3D-structure; Lipoprotein.  
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 724 COAT PROTEIN VP3.  
 FT CHAIN 725 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1107 CORE PROTEIN P12.  
 FT CHAIN 1108 1425 CORE PROTEIN P14.  
 FT CHAIN 1426 1578 CORE PROTEIN P19.  
 FT CHAIN 1579 1649 GENOME-LINKED PROTEIN VEG.  
 FT CHAIN 1650 1862 PROTEASE.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN 202 202 N-myristoyl glycine (in host).  
 FT LIPID 511 511 INTERCHAIN (IN VP3 DIMER).  
 FT DISULFID 511 511 IN VP2-VP1 DIMER.  
 FT DISULFID 406 858



```

FT VARIANT 780 780 I -> V (IN STRAIN OIBFS).
FT VARIANT 808 808 G -> R (IN STRAIN OIBFS).
FT VARIANT 861 861 N -> S (IN STRAIN OIBFS).
SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;

Query Match 75.9%; Score 30; DB 1; Length 2332;
Best Local Similarity 62.5%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 2;

Qy 2 GVGTXIRP 9
Db 114 GIGTASRP 121

RESULT 9
POLG_FMDV1 STANDARD; PRT; 2333 AA.
ID AC POLG_FMDV1 STANDARD; PRT; 2333 AA.
AC Q03306; Q84768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;
AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;
AC Q84769; Q89824;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VP1 TO
DE VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed
DE RNA polymerase P56A (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
NCBI_TaxID=12112;
RX SEQUENCE FROM N.A.
RX MEDLINE=84169547; PubMed=6324120;
RA Carroll A.R., Rowlands D.J., Clarke B.E.;
RT "The complete nucleotide sequence of the RNA coding for the primary
RL translation product of foot and mouth disease virus.";
RN Nucleic Acids Res. 12:2461-2472(1984).
RN [2]
RX SEQUENCE OF 115-1048 FROM N.A.
RX MEDLINE=82211814; PubMed=6282711;
RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;
RT "The nucleotide sequence of cDNA coding for the structural proteins
RL of foot-and-mouth disease virus.";
RN Gene 17:153-161(1982).
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC -!- {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC -----
CC EMBL; V01130; CAA24361.1; -
CC EMBL; X00429; CAA25127.1; -
CC MEROPS; C03.008; -
CC MEROPS; C28.001; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR004080; FMDVPiccoat.
CC InterPro; IPR008739; Peptidase_C28.
CC InterPro; IPR001676; Rhv.

```

```

DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_Es.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF05408; Peptidase_C28; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR PRINTS; PR01542; FMDVPiccoat.
KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;
KW Myristate; Lipoprotein.
FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.
FT CHAIN 202 286 COAT PROTEIN VP4.
FT CHAIN 287 504 COAT PROTEIN VP2.
FT CHAIN 505 725 COAT PROTEIN VP3.
FT CHAIN 726 937 COAT PROTEIN VP1.
FT CHAIN 938 1578 CORE PROTEIN P52.
FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VP1.
FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VP2.
FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VP3.
FT CHAIN 1650 1863 PROTEASE P20B.
FT CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.
FT CHAIN 202 202 N-myristoyl glycine (in host).
FT LIPID 202 202 S -> C (IN REF. 2).
FT CONFLICT 396 396 P -> L (IN REF. 2).
FT CONFLICT 632 632 P -> L (IN REF. 2).
SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 76.9%; Score 30; DB 1; Length 2333;
Best Local Similarity 62.5%; Pred. No. 2.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 2;

Qy 2 GVGTXIRP 9
Db 114 GIGTASRP 121

RESULT 10
POLG_FMDVZ STANDARD; PRT; 2336 AA.
ID AC POLG_FMDVZ STANDARD; PRT; 2336 AA.
AC P49303;
DT 01-PEB-1996 (Rel. 33, Created)
DT 01-PEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat
DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked
DE proteins VP1 TO VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].
OS Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)
OS (Aphthovirus A) (FMDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
NCBI_TaxID=73481;
OX NCBI_TaxID=73481;
RN [1]
RN SEQUENCE FROM N.A.
RA Sosnovtsev S.V., Onischenko A.M., Petrov N.A., Kalashnikova T.I.,
RA Mamaeva N.V., Drugin V.V., Perevozchikova N.A., Vasilenko S.K.;
RA Submitted (Aug-1993) to the EMBL/GenBank/DBJ databases.
RL -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC -----

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; X74812; CAA52812.1; -  
DR PIR; S37077; S37077.  
DR HSSP; Q88571; ITME.  
DR MEROPS; C03.008; -  
DR MEROPS; C28.001; -  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR004080; FMDVPicoat.  
DR InterPro; IPR008739; Peptidase\_C28.  
DR InterPro; IPR001676; Rhv.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_D5\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_P5vir.  
DR InterPro; IPR008975; Viral\_csp\_coat.  
DR Pfam; PF05408; Peptidase\_C28; 1.  
DR Pfam; PF00073; rhv; 3.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICVIRUSNS.  
DR PRINTS; PR01542; FMDVPICOAT.  
KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
KW Myristate; Lipoprotein.  
FT CHAIN 1 201  
FT CHAIN 202 286 NONSTRUCTURAL PROTEIN P20A.  
FT CHAIN 287 504 COAT PROTEIN VP4.  
FT CHAIN 505 724 COAT PROTEIN VP3.  
FT CHAIN 725 938 COAT PROTEIN VP1.  
FT CHAIN 939 954 CORE PROTEIN X.  
FT CHAIN 955 1108 CORE PROTEIN P14.  
FT CHAIN 1109 1426 CORE PROTEIN P19.  
FT CHAIN 1427 1579 CORE PROTEIN P19.  
FT CHAIN 1580 1602 GENOME-LINKED PROTEIN VP1.  
FT CHAIN 1603 1626 GENOME-LINKED PROTEIN VP2.  
FT CHAIN 1627 1650 GENOME-LINKED PROTEIN VP3.  
FT CHAIN 1651 1863 PROTEASE.  
FT CHAIN 1864 2336 RNA-DIRECTED RNA POLYMERASE.  
FT LIPID 202 202 N-myristoyl glycine (in host) (By similarity).  
SQ SEQUENCE 2336 AA; 259982 MW; 15AC2AB022B5B954 CRC64;  
Query Match 76.9%; Score 30; DB 1; Length 2336;  
Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
DB 114 GIGTASRP 121  
RESULT 11  
COGS\_UCAPU STANDARD; PRT; 226 AA.  
AC P00771.  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 15-JUL-1998 (Rel. 36, last sequence update)  
DT 28-FEB-2003 (Rel. 41, last annotation update)  
DE Brachyurin (EC 3.4.21.32) (Collagenolytic protease).  
OS Uca pugnator (Atlantic sand fiddler crab) (Celaqua pugilator).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubrachyura; Ocypodoidea; Ocypodidae; Ocypodinae; Uca complex;  
OC Celaqua.  
OX NCBI\_TaxID=6772;  
RN [1]  
RP SEQUENCE.

RC TISSUE=Hepatopancreas;  
RX MEDLINE=81040004; PubMed=6252953;  
RA Grant G.A., Henderson K.O., Eisen A.Z., Bradshaw R.A.;  
RT "Amino acid sequence of a collagenolytic protease from the  
hepatopancreas of the fiddler crab, Uca pugnator.";  
RL Biochemistry 19:4653-4659(1980).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS), AND REVISIONS.  
RC TISSUE=Hepatopancreas;  
RX MEDLINE=97299771; PubMed=9154920;  
RA Perona J.J., Tsu C.A., Craik C.S., Fletterick R.J.;  
RT "Crystal structure of an ecotin-collagenase complex suggests a model  
for recognition and cleavage of the collagen triple helix.";  
RL Biochemistry 36:5381-5392(1997).  
CC -!- FUNCTION: This enzyme is a serine protease capable of degrading  
the native triple helix of collagen.  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins, with broad specificity  
for peptide bonds. Degrades native collagen at about 75% of the  
length of the molecule from the N-terminus. Low activity on small  
molecule substrates of both trypsin and chymotrypsin.  
CC -!- SIMILARITY: Belongs to peptidase family S1.  
DR PIR; A00959; KCUF.  
DR MEROPS; S01.122; -  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR Pfam; PF00889; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPsin.  
DR SMART; SM00020; TRYp\_SPC; 1.  
DR PROSITE; PS02400; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Collagen degradation; 3D-structure.  
FT ACT\_SITE 41 41  
FT ACT\_SITE 87 87 CHARGE RELAY SYSTEM.  
FT ACT\_SITE 178 178 CHARGE RELAY SYSTEM.  
FT DISULFID 26 42  
FT DISULFID 151 164  
FT DISULFID 174 200  
FT CONFLICT 91 91 I -> V (IN REF. 1).  
FT CONFLICT 147 148 SN -> NS (IN REF. 1).  
FT CONFLICT 175 175 N -> D (IN REF. 1).  
FT CONFLICT 185 185 N -> D (IN REF. 1).  
FT STRAND 2 2  
FT STRAND 5 6  
FT TURN 9 10  
FT TURN 13 14  
FT STRAND 15 20  
FT TURN 21 23  
FT STRAND 24 32  
FT TURN 33 34  
FT STRAND 35 38  
FT HELIX 40 43  
FT TURN 44 45  
FT STRAND 49 53  
FT STRAND 57 57  
FT TURN 63 64  
FT STRAND 66 70  
FT STRAND 73 75  
FT TURN 77 78  
FT STRAND 80 80  
FT TURN 81 84  
FT STRAND 85 85  
FT STRAND 89 92  
FT STRAND 107 107  
FT TURN 115 116  
FT STRAND 118 123  
FT TURN 129 130  
FT STRAND 137 137  
FT STRAND 139 140  
FT STRAND 143 146  
FT HELIX 148 155

FT TURN 160 161  
FT STRAND 162 165  
FT TURN 168 170  
FT STRAND 172 172  
FT TURN 175 176  
FT TURN 178 179  
FT STRAND 181 184  
FT TURN 185 186  
FT STRAND 187 196  
FT TURN 197 198  
FT TURN 200 201  
FT STRAND 206 210  
FT HELIX 212 214  
FT TURN 215 222  
SQ SEQUENCE 226 AA; E45591CAF332CE8C CRC64;  
  
Query Match 74.4%; Score 29; DB 1; Length 226;  
Best Local Similarity 62.5%; Pred. No. 46;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVGTXIRP 9  
Db 114 GVGTXIRP 121  
  
RESULT 12  
TS1 STRCS  
ID T251 STRCS STANDARD; PRT; 227 AA.  
AC Q52691;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Type II restriction enzyme ScaI (EC 3.1.21.4) (Endonuclease ScaI)  
DE (R.ScaI).  
GN SCAIR.  
OS Streptomyces caespitosus.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomyces; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=53502;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99077292; PubMed=9862476;  
RA Xu S.-Y., Xiao J.-P., Ettwiller L., Holden M., Aliotta J., Poh C.L.,  
RA Dalton M., Robinson D.P., Petronzio T.R., Moran L., Ganatra M.,  
RA Ware J., Slatko B., Benner J. II;  
RA "Cloning and expression of the ApsLI, Napi, NaphI, ScaI, ScaI, and  
RT Sapi restriction-modification systems in Escherichia coli."  
RL Mol. Gen. Genet. 260:226-231(1998).  
CC -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE AGTACT AND  
CC CLEAVAGES AFTER T-3.  
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  
CC specific double-stranded fragments with terminal 5'-phosphates.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF044681; AAC97177.1; -  
DR REBASE; 1628; ScaI.  
KW Restriction system; Hydrolase; Nuclease; Endonuclease.  
SQ SEQUENCE 227 AA; 25463 MW; A4E4807275B6009 CRC64;  
  
Query Match 74.4%; Score 29; DB 1; Length 227;  
Best Local Similarity 71.4%; Pred. No. 47;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 GVGTXIRP 9  
Db 81 IGTQIRP 87

RESULT 13  
ARGB\_CHRVO  
ID ARGB\_CHRVO STANDARD; PRT; 289 AA.  
AC Q7NR63;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK) (N-acetyl-L-  
DE glutamate 5-phosphotransferase).  
DE ARGB OR CV3921.  
GN Chromobacterium violaceum.  
OS Chromobacterium violaceum.  
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
OC Neisseriaceae; Chromobacterium.  
OX NCBI\_TaxID=536;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 12472 / DSM 30191;  
RX MEDLINE=22882880; PubMed=14500782;  
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimarães C.T.,  
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,  
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,  
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,  
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,  
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,  
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,  
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,  
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,  
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,  
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furian L.R.,  
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,  
RA Grattapaglia D., Grigard E.C., Hanna E.S., Jardim S.N., Laurino J.,  
RA Leol L.C.T., Lima L.F.A., Loureiro M.F., Lyza M.C.C.P.,  
RA Madeira H.M.F., Manlio G.P., Maranhao A.Q., Martins W.S.,  
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,  
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,  
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,  
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,  
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,  
RA Santos E.B.P., Santos P.R., Schneider M.P.C., Seunarez H.N.,  
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,  
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,  
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,  
RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;  
RT "The complete genome sequence of Chromobacterium violaceum reveals  
RT remarkable and exploitable bacterial adaptability."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:11650-11655(2003).  
CC -!- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-  
CC glutamate 5-phosphate.  
CC -!- PATHWAY: Arginine biosynthesis; second step.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the acetylglutamate kinase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF016923; AAQ61582.1; -  
DR HAMAP; MF\_00082; -; 1.  
KW Arginine Biosynthesis; Transferase; Kinase; Complete proteome.  
FT ACT\_SITE 30 30  
FT SITE 85 95 BY SIMILARITY.  
FT SITE 87 87 SUBSTRATE BINDING (BY SIMILARITY).  
FT SITE 187 187 SUBSTRATE BINDING (BY SIMILARITY).  
FT ACT\_SITE 247 247 BY SIMILARITY.  
FT SEQUENCE 289 AA; 30353 MW; D68B450ED63F92B CRC64;  
  
Query Match 74.4%; Score 29; DB 1; Length 289;

Best Local Similarity 85.7%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIR 8  
Db 282 GVGTMIR 288

RESULT 14

ARGB\_ARCFU STANDARD; PRT; 290 AA.  
AC O28986;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Probable acetylglutamate kinase (EC 2.7.2.8) (NAG kinase) (AGK)  
DE (N-acetyl)-L-glutamate 5-phosphotransferase).  
GN ARGB OR AF1280.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirschner S.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,  
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370 (1997).  
CC -!- CATALYTIC ACTIVITY: ATP + N-acetyl-L-glutamate = ADP + N-acetyl-L-  
CC glutamate 5-phosphate.  
CC -!- PATHWAY: Arginine biosynthesis; second step.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- SIMILARITY: Belongs to the acetylglutamate kinase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AE001016; AAB89966.1; -  
CC PIR; G69409; G69409.  
CC TIGR; AF1280; -  
CC HAMAP; MF\_00082; -; 1.  
CC InterPro; IPR001048; Aa\_kinase.  
CC DR InterPro; IPR004662; Acglut\_kinase.  
CC DR Pfam; PF00696; aak\_kinase; 1.  
CC DR TIGRFAMs; TIGR00761; argB; 1.  
KW Arginine biosynthesis; Transferase; Kinase; Complete proteome.  
FT ACT\_SITE 25 25 BY SIMILARITY.  
FT ACT\_SITE 248 248 BY SIMILARITY.  
FT SITE 80 80 SUBSTRATE BINDING (BY SIMILARITY).  
FT SITE 82 80 SUBSTRATE BINDING (BY SIMILARITY).  
FT SITE 185 185 SUBSTRATE BINDING (BY SIMILARITY).  
SQ SEQUENCE 290 AA; 31664 MW; 08205961A35EF37C CRC64;

Query Match 74.4%; Score 29; DB 1; Length 290;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 283 GIGTWEP 290

RESULT 15

MPG1\_YEAST STANDARD; PRT; 361 AA.  
AC P41940;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) (ATP-mannose-1-  
DE phosphate guanylyltransferase) (NDP-hexose pyrophosphorylase).  
GN MPG1 OR PSA1 OR YDL055C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96418871; PubMed=8921656;  
RA Benton B.K., Plump S.D., Roos J., Lennarz W.J., Cross F.R.;  
RT "Over-expression of S. cerevisiae G1 cyclins restores the viability  
RT of alg1 N-glycosylation mutants.";  
RL Curr. Genet. 29:106-113 (1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Schultz J., Sprague G.F. Jr.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Bloescher H., Brandt P.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Essential protein which may be involved in the  
CC regulation of cell cycle progression.  
CC -!- CATALYTIC ACTIVITY: GTP + alpha-D-mannose 1-phosphate =  
CC diphosphate + GDP-mannose  
CC -!- SIMILARITY: TO S-TYRPHIMORIUM CDP-GLUCOSE PYROPHOSPHYLASE (RFBP).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U19608; AAC49289.1; -  
CC EMBL; U24437; AAA69677.1; -  
CC EMBL; Z74103; CAA98617.1; -  
CC PIR; S67590; S67590.  
CC Germonline; 140297; -  
CC SGD; S0002213; PSA1.  
CC DR GO; GO:0005737; Cytoplasm; IDA.  
CC DR GO; GO:0004475; F:mannose-1-phosphate guanylyltransferase act. .; IDA.  
CC DR GO; GO:0000032; P:cell wall mannoprotein biosynthesis; IMP.  
CC DR GO; GO:0003298; P:GDP-mannose biosynthesis; IDA.  
CC DR GO; GO:0006486; P:protein amino acid glycosylation; IMP.  
CC InterPro; IPR005835; NTP\_transferase.  
CC DR Pfam; PF00132; hexapep; 4.  
CC DR Pfam; PF00483; NTP\_transferase; 1.  
CC DR PROSITE; PS00101; HEXAPEP TRANSFERASES; UNKNOWN\_2.  
CC FT TRANSFERASE; Kinase; Cell-cycle.  
CC FT CONFLICT 50 50 V -> A (IN REF. 2).  
SQ SEQUENCE 361 AA; 39566 MW; DBF1C39BEAE0B776 CRC64;

Query Match 74.4%; Score 29; DB 1; Length 361;  
Best Local Similarity 62.5%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9

Db           | | | : | |  
              9 GYGRLRP 16

Search completed: April 1, 2004, 17:39:25  
Job time : 6.52632 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 28.4211 Seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-3  
Perfect score: 29  
Sequence: 1 XGVGTIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacterioplasm:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	89.7	53	17 Q9HS92	Q9HS92 halobacteri
2	34	87.2	256	17 Q8U0F4	Q8U0F4 pyrococcus
3	34	87.2	2330	12 Q7TDB3	Q7TDB3 foot-and-mo
4	32	82.1	83	16 Q8DM12	Q8DM12 streptococc
5	32	82.1	203	12 Q8JVD4	Q8JVD4 foot-and-mo
6	32	82.1	203	12 Q8JVD2	Q8JVD2 foot-and-mo
7	32	82.1	203	12 Q8JVE0	Q8JVE0 foot-and-mo
8	32	82.1	203	12 Q8JVD5	Q8JVD5 foot-and-mo
9	32	82.1	203	12 Q8JVD6	Q8JVD6 foot-and-mo
10	32	82.1	203	12 Q8JVD3	Q8JVD3 foot-and-mo
11	32	82.1	203	12 Q8JVD9	Q8JVD9 foot-and-mo
12	32	82.1	203	12 Q8JUW5	Q8JUW5 foot-and-mo
13	32	82.1	203	12 Q8JVD8	Q8JVD8 foot-and-mo
14	32	82.1	305	16 Q98K87	Q98K87 rhizobium l
15	32	82.1	347	16 Q92259	Q92259 rhizobium m
16	32	82.1	396	16 Q8DB02	Q8DB02 vibrio vuln

17	32	82.1	462	16 P75023	P75023 synchocyst
18	32	82.1	1291	13 Q8DGN7	Q8DGN7 lampetra fl
19	31	79.5	152	12 Q8JSZ5	Q8JSZ5 prunus necr
20	31	79.5	203	12 Q8JVD1	Q8JVD1 foot-and-mo
21	31	79.5	205	12 Q8JVE4	Q8JVE4 foot-and-mo
22	31	79.5	225	17 Q8ZYC7	Q8ZYC7 pyrobaculum
23	31	79.5	226	12 Q9IMS3	Q9IMS3 prunus necr
24	31	79.5	226	12 Q9YKES	Q9YKES prunus necr
25	31	79.5	226	12 Q98600	Q98600 prunus necr
26	31	79.5	226	12 Q9ICF4	Q9ICF4 prunus necr
27	31	79.5	226	12 Q9IMS0	Q9IMS0 prunus necr
28	31	79.5	227	16 Q8XMQ9	Q8XMQ9 clostridium
29	31	79.5	229	2 Q93M14	Q93M14 streptococc
30	31	79.5	229	16 Q97QE9	Q97QE9 streptococc
31	31	79.5	229	16 Q8DF16	Q8DF16 streptococc
32	31	79.5	240	16 Q988F3	Q988F3 rhizobium l
33	31	79.5	243	16 Q92TC3	Q92TC3 rhizobium l
34	31	79.5	281	16 Q839C5	Q839C5 enterococcu
35	31	79.5	342	2 Q9KIU4	Q9KIU4 myxococcus
36	31	79.5	349	16 Q8RDG7	Q8RDG7 thermoanaer
37	31	79.5	353	2 Q9ZA42	Q9ZA42 streptomyce
38	31	79.5	361	16 Q9RZC3	Q9RZC3 deinococcus
39	31	79.5	364	16 Q88DJ4	Q88DJ4 pseudomonas
40	31	79.5	366	12 Q91TT6	Q91TT6 tupalaia herp
41	31	79.5	366	16 Q9H2T9	Q9H2T9 pseudomonas
42	31	79.5	377	16 Q9PFS8	Q9PFS8 ureaplasma
43	31	79.5	430	2 Q931E6	Q931E6 propionibac
44	31	79.5	437	16 Q9PFA1	Q9PFA1 xylalla fas
45	31	79.5	437	16 Q87BE0	Q87BE0 xylalla fas

## ALIGNMENTS

### RESULT 1

Q9HS92	PRELIMINARY;	PRT;	53 AA.
ID Q9HS92			
AC Q9HS92			
DT 01-MAR-2001	(TRENBLrel. 16, Created)		
DT 01-MAR-2001	(TRENBLrel. 16, Last sequence update)		
DT 01-JUN-2003	(TRENBLrel. 24, Last annotation update)		
DE VNG0346h.			
GN VNG0346h.			
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).			
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
OC Halobacteriaceae; Halobacterium.			
OX NCBI_TaxID=64091;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE=2050483; PubMed=11016950;			
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,			
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,			
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,			
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,			
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,			
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,			
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,			
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;			
RT "Genome sequence of Halobacterium species NRC-1.";			
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).			
DR EMEL; AE004994; AAG18916.1; -.			
DR PIR; H84193; H84193.			
KW Complete proteome.			
SQ SEQUENCE 53 AA; 5115 MW; 759B14B1CD128573 CRC64;			

Query Match 89.7%; Score 35; DB 17; Length 53;  
Best Local Similarity 75.0%; Pred. No. 3.4;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9

DB 34 GVGTXIRP 41

```

RESULT 2
Q8U0F4 PRELIMINARY; PRT; 256 AA.
ID Q8U0F4
AC Q8U0F4; 2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Acetylglutamate kinase, putative.
GN P1636.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010263; AAJ81760.1; -.
DR GO; GO:0016301; P:kinase activity; IEA.
DR InterPro; IPR001048; Aa_kinase.
DR Pfam; PF00696; aak_kinase; 1.
DR Kinase; Complete proteome.
KW Kinase; Complete proteome.
SQ SEQUENCE 256 AA; 28443 MW; 7D88B6B3E1F0676F CRC64;

Query Match 87.2%; Score 34; DB 17; Length 256;
Best Local Similarity 62.5%; Pred. No. 29;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9
|:|:|:|
DB 249 GIGTIVRP 256

RESULT 3
Q7TDB3 PRELIMINARY; PRT; 2330 AA.
ID Q7TDB3
AC Q7TDB3; 2002 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Polyprotein.
OS Foot-and-mouth disease virus Asia 1.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=110195;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=vaccine IND 63/72;
RA Saravanan T., Reddy G.R., Dechamma H.J., Suryanarayana V.V.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY304994; AAP60035.1; -.
KW Polyprotein.
FT CHAIN 1 201 L-protease.
FT CHAIN 202 951 P1-2A.
FT CHAIN 952 1105 2B.
FT CHAIN 1106 1423 2C.
FT CHAIN 1424 1647 3AB.
FT CHAIN 1648 1860 3C.
FT CHAIN 1861 2330 3D.
SQ SEQUENCE 2330 AA; 258793 MW; 8135278D750BDB02 CRC64;

Query Match 87.2%; Score 34; DB 12; Length 2330;
Best Local Similarity 62.5%; Pred. No. 2.9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9
|:|:|:|
DB 114 GIGTALRP 121

RESULT 4
Q8DW12 PRELIMINARY; PRT; 83 AA.
ID Q8DW12
AC Q8DW12; 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN SMU.277.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
RA MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Perretti J.J.;
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
pathogen.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014876; AAN58044.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 83 AA; 8140 MW; 65EF42894C95B6A6 CRC64;

Query Match 82.1%; Score 32; DB 16; Length 83;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9
|:|:|:|
DB 46 GVGTAIEP 53

RESULT 5
Q8JVD4 PRELIMINARY; PRT; 203 AA.
ID Q8JVD4
AC Q8JVD4; 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Polyprotein (Fragment).
OS Foot-and-mouth disease virus SAT 2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Aphthovirus.
OX NCBI_TaxID=35292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAT 2;
RX MEDLINE=22032953; PubMed=12036580;
RA van Rensburg H., Haydon D., Joubert F., Bastos A., Heath L., Nel L.;
RT "Genetic heterogeneity in the foot-and-mouth disease virus leader and
3C proteinases.";
RL Gene 289:19-29(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SAT 2;
RA Van Rensburg H.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF283442; AAM48507.1; -.
DR InterPro; IPR008739; Peptidase_C28.
DR Pfam; PF05408; Peptidase_C28; 1.
DR CHAIN 1 >199 LEADER PROTEINASE.
FT CHAIN 200 >203 VF4.
FT CHAIN 203 203
FT NON_TER 203
SQ SEQUENCE 203 AA; 23377 MW; D1AE0F9B3D6CC280 CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 6  
Q8JVD2 PRELIMINARY; PRT; 203 AA.  
ID Q8JVD2  
AC Q8JVD2; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 3.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=12123;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 3;  
RX MEDLINE=22032953; PubMed=12036580;  
RA van Rensburg H.G.; Haydon D., Joubert F., Bastos A., Heath L., Nel L.;  
RT "Genetic heterogeneity in the foot-and-mouth disease virus Leader and  
3C proteinases";  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON TER 203  
SQ SEQUENCE 203 AA; 2326 MW; F92340A598419C4E CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 7  
Q8JVE0 PRELIMINARY; PRT; 203 AA.  
ID Q8JVE0  
AC Q8JVE0; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 1.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=12122;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 1;  
RX MEDLINE=22032953; PubMed=12036580;  
RA van Rensburg H.G.; Haydon D., Joubert F., Bastos A., Heath L., Nel L.;  
RT "Genetic heterogeneity in the foot-and-mouth disease virus Leader and  
3C proteinases";  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON TER 203  
SQ SEQUENCE 203 AA; 2326 MW; F92340A598419C4E CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 8  
Q8JVD5 PRELIMINARY; PRT; 203 AA.  
ID Q8JVD5  
AC Q8JVD5; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=35292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 2;  
RX MEDLINE=22032953; PubMed=12036580;  
RA van Rensburg H.G.; Haydon D., Joubert F., Bastos A., Heath L., Nel L.;  
RT "Genetic heterogeneity in the foot-and-mouth disease virus Leader and  
3C proteinases";  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON TER 203  
SQ SEQUENCE 203 AA; 23205 MW; DE1A145DDE8B3588 CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 9  
Q8JVD6 PRELIMINARY; PRT; 203 AA.  
ID Q8JVD6  
AC Q8JVD6; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=35292;  
RN [1]  
RP SEQUENCE FROM N.A.

DR EMBL; AF283436; AAM48501.1; -.  
DR InterPro; IPR008739; Peptidase\_C28.  
DR Pfam; PF05408; Peptidase\_C28; 1.  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON TER 203  
SQ SEQUENCE 203 AA; 23243 MW; 93F912C3A61D8595 CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 8  
Q8JVD5 PRELIMINARY; PRT; 203 AA.  
ID Q8JVD5  
AC Q8JVD5; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=35292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 2;  
RX MEDLINE=22032953; PubMed=12036580;  
RA van Rensburg H.G.; Haydon D., Joubert F., Bastos A., Heath L., Nel L.;  
RT "Genetic heterogeneity in the foot-and-mouth disease virus Leader and  
3C proteinases";  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON TER 203  
SQ SEQUENCE 203 AA; 23205 MW; DE1A145DDE8B3588 CRC64;

Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVGTXIRP 9  
Db 112 GVGSTSRP 119

RESULT 9  
Q8JVD6 PRELIMINARY; PRT; 203 AA.  
ID Q8JVD6  
AC Q8JVD6; 82.1%; Score 32; DB 12; Length 203;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 2.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=35292;  
RN [1]  
RP SEQUENCE FROM N.A.



```

Query Match      82.1%; Score 32; DB 12; Length 203;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      2  GVGTXRTP 9
      |||||
Db      112  GVGTSTRP 119

```

DR InterPro; IPR003171; Methylhydrof\_redctase.  
DR Pfam; PF02219; MTHFR; 1.  
DR TIGRFAMs; TIGR00676; fadh2; 1.  
KW Complete proteome.  
SQ SEQUENCE 305 AA; 33577 MW; 7658906AB702ACB8 CRC64;  
Query Match 82.1%; Score 32; DB 16; Length 305;  
Best Local Similarity 75.0%; Pred. No. 97;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
DB 117 GVGAYRVP 124  
RESULT 15  
Q92Z59 PRELIMINARY; PRT; 347 AA.  
AC Q92Z59;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative hydrolase protein.  
DE RA0633 OR Sma1166.  
GN Rhizobium meliloti (Sinorhizobium meliloti).  
OS Plasmid pSymA (megaplasmid 1).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396509; PubMed=11481432;  
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abolia A.P.,  
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,  
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn M.J.,  
RA Kallman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,  
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;  
RT "Nucleotide sequence and predicted functions of the entire  
RT Sinorhizobium meliloti pSymA megaplasmid."  
RT Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).  
DR EMBL; AF007253; AAK65291.1; -.  
DR PIR; A95341; A95341.  
DR GO; GO:0045821; C:extrachromosomal DNA; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.  
DR InterPro; IPR000073; A/b hydrolase.  
DR InterPro; IPR003089; AB\_Hydrolase.  
DR InterPro; IPR000639; Epox\_Hydrolase.  
DR InterPro; IPR000379; Ser\_estrs.  
DR Pfam; PF00561; abhydrolase; 1.  
DR PRINTS; PR00111; ABHYDROLASE.  
DR PRINTS; PR00412; EPOXYHYDROLASE.  
DR Hydroxylase; Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 347 AA; 35861 MW; F802934A261686E CRC64;  
Query Match 82.1%; Score 32; DB 16; Length 347;  
Best Local Similarity 62.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
DB 328 GAGTSVRP 335  
Search completed: April 1, 2004, 17:38:40  
Job time : 31.4211 secs

RESULT 13  
Q8JVD8 PRELIMINARY; PRT; 203 AA.  
AC Q8JVD8;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Polyprotein (Fragment).  
OS Foot-and-mouth disease virus SAT 1.  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OX NCBI\_TaxID=12122;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 1;  
RX MEDLINE=22032953; PubMed=12036580;  
RA van Rensburg H., Haydon D., Joubert F., Bastos A., Heath L., Nel L.;  
RT "Genetic heterogeneity in the foot-and-mouth disease virus Leader and  
RT 3C proteinases";  
RL Gene 289:19-29(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SAT 1;  
RX Van Rensburg H.G.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF283438; AAM48503.1; -.  
DR InterPro; IPR008739; Peptidase\_C28.  
DR Pfam; PF05408; Peptidase\_C28; 1.  
FT CHAIN 1 >199 LEADER PROTEINASE.  
FT CHAIN 200 >203 VP4.  
FT NON\_TER 203  
FT SEQUENCE 203 AA; 23152 MW; 0C555FEA7BFC7987 CRC64;  
Query Match 82.1%; Score 32; DB 12; Length 203;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVGTXIRP 9  
DB 112 GVGSTRP 119  
RESULT 14  
Q98K87 PRELIMINARY; PRT; 305 AA.  
AC Q98K87;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 5-10-methylenetetrahydrofolate reductase.  
GN M11587.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP002997; BAB48927.1; -.  
DR GO; GO:0008702; F:5,10-methylenetetrahydrofolate reductase (F. . .; IEA.  
DR GO; GO:0004489; F:methylenetetrahydrofolate reductase (NADPH) . . .; IEA.  
DR GO; GO:0009086; P:methionine biosynthesis; IEA.  
DR InterPro; IPR004620; Fadh2\_bact.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 43.1053 Seconds  
(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-4

Perfect score: 34  
Sequence: 1 XGVXVIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Genesep 29Jan04:\*  
1: genesep1980s.\*  
2: genesep1990s.\*  
3: genesep2000s.\*  
4: genesep2001s.\*  
5: genesep2002s.\*  
6: genesep2003as.\*  
7: genesep2003bs.\*  
8: genesep2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	88.2	8	7 ADD31284	Add31284 Angiogene
2	30	88.2	8	7 ADE83635	Ade83635 Antiangi
3	30	88.2	9	3 AAY67983	Aay67983 Antiangi
4	30	88.2	9	4 AAU15042	Aau15042 Antiangi
5	30	88.2	9	4 AAU15043	Aau15043 Antiangi
6	30	88.2	9	4 AAU15027	Aau15027 Antiangi
7	30	88.2	9	6 AAO26633	Aao26633 Anti-angi
8	30	88.2	10	4 AAU4987	Aau4987 Antiangi
9	30	88.2	10	4 AAU14993	Aau14993 Antiangi
10	30	88.2	139	6 ABP57616	Abp57616 S. muraya
11	30	88.2	2796	6 ABU79133	Abu79133 Mycobacte
12	29	85.3	8	7 ADD31274	Add31274 Angiogene
13	29	85.3	8	7 ADD31265	Add31265 Angiogene
14	29	85.3	8	7 ADD31264	Add31264 Angiogene
15	29	85.3	8	7 ADD31266	Add31266 Angiogene
16	29	85.3	8	7 ADD31294	Add31294 Angiogene
17	29	85.3	8	7 ADD31268	Add31268 Angiogene
18	29	85.3	8	7 ADD31273	Add31273 Angiogene
19	29	85.3	8	7 ADD31269	Add31269 Angiogene
20	29	85.3	8	7 ADD31267	Add31267 Angiogene
21	29	85.3	8	7 ADD31282	Add31282 Angiogene
22	29	85.3	8	7 ADE83640	Ade83640 Antiangi
23	29	85.3	8	7 ADE83694	Ade83694 Antiangi
24	29	85.3	8	7 ADE83632	Ade83632 Antiangi
25	29	85.3	8	7 ADE83696	Ade83696 Antiangi

26	29	85.3	8	7 ADE83629	Ade83629 Antiangi
27	29	85.3	8	7 ADE83628	Ade83628 Antiangi
28	29	85.3	8	7 ADE83641	Ade83641 Antiangi
29	29	85.3	8	7 ADE83631	Ade83631 Antiangi
30	29	85.3	8	7 ADE83691	Ade83691 Antiangi
31	29	85.3	9	3 AAY67973	Aay67973 Antiangi
32	29	85.3	9	3 AAY67978	Aay67978 Antiangi
33	29	85.3	9	3 AAY67992	Aay67992 Antiangi
34	29	85.3	9	3 AAY67975	Aay67975 Antiangi
35	29	85.3	9	3 AAY67986	Aay67986 Antiangi
36	29	85.3	9	3 AAY67996	Aay67996 Antiangi
37	29	85.3	9	3 AAY67997	Aay67997 Antiangi
38	29	85.3	9	3 AAY67979	Aay67979 Antiangi
39	29	85.3	9	3 AAY67995	Aay67995 Antiangi
40	29	85.3	9	3 AAY67985	Aay67985 Antiangi
41	29	85.3	9	3 AAY67976	Aay67976 Antiangi
42	29	85.3	9	3 AAY67989	Aay67989 Antiangi
43	29	85.3	9	3 AAY67974	Aay67974 Antiangi
44	29	85.3	9	3 AAY67981	Aay67981 Antiangi
45	29	85.3	9	3 AAY67982	Aay67982 Antiangi

## ALIGNMENTS

RESULT 1  
ADD31284  
ID ADD31284 standard; peptide; 8 AA.  
XX  
AC ADD31284;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Angiogenesis inhibiting peptide #59.  
XX  
KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX  
OS Synthetic.  
XX

Key Key Location/Qualifiers  
FT Modified-site 1 /label= OTHER  
FT /note= "OTHER= N-acetyl"  
FT Modified-site 3 /label= OTHER  
FT /note= "OTHER= D-form residue, alle (not defined)"  
FT Modified-site 8 /note= "NHCH2CH3"  
XX US2003109455-A1.  
XX 12-JUN-2003.  
XX 30-OCT-2002; 2002US-00283550.  
XX 31-OCT-2001; 2001US-0335017P.  
XX (HAYV/) HAYV F.  
XX (BRAD/) BRADLEY M F.  
XX Haviv F, Bradley MF;  
XX WPI; 2003-843101/78.  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
PS Claim 13; Page 25; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX SQ Sequence 8 AA;  
 Query Match 88.2%; Score 30; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 GVXTXIRP 9  
 Db 1 GVXTSIRP 8  
 |||||  
 |||||

RESULT 2  
 ADE83635  
 ID ADE83635 standard; peptide; 8 AA.  
 XX AC ADE83635;  
 XX DT 29-JAN-2004 (first entry)  
 XX DE Antiangiogenic peptide.

XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 XX antiarthritic; ophthalmological; antidiabetic; dermatological;  
 XX KW antipsoriatic; antitumor; antidiabetic; cardiant; vulnary;  
 KW antiinflammatory; antitumor; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.  
 XX KW Synthetic.

XX PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "acetylated"  
 FT Modified-site 3  
 FT Modified-site 8 /note= "D-form residue alloisoleucyl"  
 FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"  
 XX WO2003037268-A2.  
 XX 08-MAY-2003.  
 XX 30-OCT-2002; 2002WO-US034811.  
 XX 31-OCT-2001; 2001US-0000681.  
 XX 04-OCT-2002; 2002US-00263812.  
 XX (ABBO ) ABBOTT LAB.  
 XX Haviv F, Bradley MF;  
 XX WPI; 2003-617886/58.  
 XX New hepta-, octa- and nona-peptide compounds used for treating e.g.  
 FT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic  
 FT retinopathy.  
 XX Claim 13; Page 46; 51pp; English.

CC The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antidiabetic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antitumor, antitumor, antitumor,  
 CC cardiant, vulnary, antinflammatory, antitumor and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumours. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilia, joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.

XX SQ Sequence 8 AA;  
 Query Match 88.2%; Score 30; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 GVXTXIRP 9  
 Db 1 GVXTSIRP 8  
 |||||  
 |||||

RESULT 3  
 RAY67983  
 ID AAY67983 standard; peptide; 9 AA.  
 XX AC AAY67983;  
 XX DT 11-APR-2000 (first entry)  
 XX DE Antiangiogenic peptide #12.

XX antiangiogenic; antiarthritic; cytostatic; antipsoriatic; arthritis;  
 KW angiogenesis; inhibitor; cancer; psoriasis; infection; ocular disease;  
 KW macular degeneration; diabetic retinopathy; tumour metastasis;  
 KW autoimmune disease; neovascularisation; Crohn's disease; birth control;  
 KW cat scratch disease.

XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /label= MeGly  
 FT /note= "Sarcosine (methylglycine); acetylated"  
 FT Misc-difference 4 /note= "D form residue"  
 FT Modified-site 9 /note= "Pro is modified to ProNHCH2CH3"  
 XX WO9961476-A1.  
 XX 02-DEC-1999.  
 XX 21-MAY-1999; 99WO-US011448.  
 XX 22-MAY-1998; 98US-00083745.  
 XX 16-FEB-1999; 99US-00250574.  
 XX 26-MAR-1999; 99US-00277466.  
 XX (ABBO ) ABBOTT LAB.  
 XX Henkin J, Haviv F, Bradley MF, Kalvin DM, Schneider AJ;  
 XX WPI; 2000-072606/06.

FT New anti-angiogenic peptides, used for treating e.g. cancer, arthritis,  
 PT psoriasis, or angiogenesis of the eye associated with infection or  
 PT surgical intervention, macular degeneration and diabetic retinopathy.  
 XX  
 PS Claim 12; Page 75; 223pp; English.

XX AAY67972 to AAY67999 represent novel anti-angiogenic peptides. Peptide  
 CC from the present invention have the formula: A0-A1-A2-A3-A4-A5-A6-A7-A8-  
 CC A9-A10 where A0 is selected from hydrogen or an acyl group; A10 is a  
 CC hydroxyl group or an amino acid amide; and A1 to A9 are amino acyl  
 CC residues. The peptides are used for anti-angiogenesis therapy in  
 CC patients. They are used to treat e.g. cancer, arthritis, psoriasis,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention, macular degeneration and diabetic retinopathy. They can  
 CC prevent tumour metastases. Further uses include treatment and prophylaxis  
 CC of autoimmune diseases, various ocular diseases e.g. diabetic  
 CC retinopathy, and other abnormal neovascularisation conditions of the eye,  
 CC skin diseases e.g. psoriasis, diseases characterised by excessive or  
 CC abnormal stimulation of endothelial cells, e.g. Crohn's disease. They can  
 CC also be used as a birth control agent, inhibiting ovulation and placental  
 CC establishment, and to treat diseases that have angiogenesis as a  
 CC pathological consequence e.g. cat scratch disease and are useful to  
 CC reduce bleeding by administration prior to surgery, especially for the  
 CC treatment of tumours. They can also be used to isolate a receptor from an  
 CC endothelial cell  
 XX  
 SQ Sequence 9 AA;

Query Match 88.2%; Score 30; DB 3; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 || || || || ||  
 DB 2 GVITAIRP 9

RESULT 4  
 AAU15042  
 ID AAU15042 standard; peptide; 9 AA.

AC AAU15042;  
 XX  
 DT 04-DEC-2001 (first entry)

XX Antiangiogenic peptide #69 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Modified-site 4  
 FT /note= "D-form residue, alioisoleucine"  
 FT Modified-site 6  
 FT /note= "N-Methyl serine"  
 FT Modified-site 9  
 FT /note= "Modified by NH-ethyl"

XX WO200138397-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032105.

XX 22-NOV-1999; 99US-00447099.

PR 31-OCT-2000; 2000US-00702649.  
 XX  
 PA (ABBO ) ABBOTT LAB.

PI Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.

XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 PT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 PT diabetic retinopathy.  
 XX  
 XX Claim 39; Page 89; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers, e.g.  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases, e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #69  
 XX  
 SQ Sequence 9 AA;

Query Match 88.2%; Score 30; DB 4; Length 9;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 || || || || ||  
 DB 2 GVXTSIRP 9

RESULT 5  
 AAU15043  
 ID AAU15043 standard; peptide; 9 AA.

XX AAU15043;

XX 04-DEC-2001 (first entry)

XX Antiangiogenic peptide #70 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Misc-difference 4  
 FT /note= "D-form residue"  
 FT Modified-site 6  
 FT /note= "N-methyl serine"  
 FT Modified-site 9

```

FT XX /note= "Modified by NH-ethyl"
PN W0200138397-A1.
XX 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US032105.
XX 22-NOV-1999; 99US-00447099.
PR 31-OCT-2000; 2000US-00702649.
XX (ABBO ) ABBOTT LAB.
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;
XX WPI; 2001-521804/57.
XX New N-alkylated peptide compounds useful for treating e.g. cancer,
PT autoimmune diseases, arthritis, psoriasis, macular degeneration and
PT diabetic retinopathy.
XX Claim 39; Page 89; 95pp; English.
XX The present invention relates to novel synthetic antiangiogenic peptides
CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The
CC antiangiogenic peptides are useful for isolating a receptor from an
CC endothelial cell. The peptides of the invention are also useful for
CC treating cancer, arthritis, psoriasis and other skin diseases,
CC angiogenesis of the eye associated with infection or surgical
CC intervention and other ocular diseases, cat scratch disease, ulcers,
CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.
CC haemangiomas and capillary action within atherosclerotic plaques,
CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,
CC Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiodysplasia,
CC wound granulation, excessive or abnormal stimulation of endothelial
CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,
CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic
CC peptides are also useful as birth control agents by inhibiting ovulation
CC and to reduce bleeding by administration before surgery. The peptides of
CC the invention exhibit improved metabolic stability, improved
CC pharmacokinetics, increased water solubility, and improved oral
CC availability. The present sequence represents antiangiogenic peptide #70
XX Sequence 9 AA;
XX Query Match 88.2%; Score 30; DB 4; Length 9;
XX Best Local Similarity 75.0%; Pred. No. 1.4e+06;
XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 GVXTXIRP 9
Db 2 GVITSIRP 9
XX 2 GVXTXIRP 9
XX 2 GVITSIRP 9
XX 04-DEC-2001 (first entry)
XX Antiangiogenic peptide #54 useful for inhibiting angiogenesis.
XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;
XX arthritis; skin disease; ocular disease; diabetic retinopathy;
XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;
XX cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.
XX Synthetic.
XX Key Location/Qualifiers
FH

```

```

FT XX /label= MeGly
FT /note= "N-methyl-glycine (sarcosine), additionally
FT modified by N-terminal acetyl"
FT Misc-difference 4 /note= "D-form residue"
FT Modified-site 6 /note= "N-Methyl alanine"
FT Modified-site 9 /note= "Modified by NH-ethyl"
XX W0200138397-A1.
XX 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US032105.
XX 22-NOV-1999; 99US-00447099.
PR 31-OCT-2000; 2000US-00702649.
XX (ABBO ) ABBOTT LAB.
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;
XX WPI; 2001-521804/57.
XX New N-alkylated peptide compounds useful for treating e.g. cancer,
PT autoimmune diseases, arthritis, psoriasis, macular degeneration and
PT diabetic retinopathy.
XX Claim 39; Page 89; 95pp; English.
XX The present invention relates to novel synthetic antiangiogenic peptides
CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The
CC antiangiogenic peptides are useful for isolating a receptor from an
CC endothelial cell. The peptides of the invention are also useful for
CC treating cancer, arthritis, psoriasis and other skin diseases,
CC angiogenesis of the eye associated with infection or surgical
CC intervention and other ocular diseases, cat scratch disease, ulcers,
CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.
CC haemangiomas and capillary action within atherosclerotic plaques,
CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,
CC Osler-Webber syndrome, myocardial angiogenesis, plaque
CC neovascularisation, telangiectasia, haemophilic joints, angiodysplasia,
CC wound granulation, excessive or abnormal stimulation of endothelial
CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,
CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic
CC peptides are also useful as birth control agents by inhibiting ovulation
CC and to reduce bleeding by administration before surgery. The peptides of
CC the invention exhibit improved metabolic stability, improved
CC pharmacokinetics, increased water solubility, and improved oral
CC availability. The present sequence represents antiangiogenic peptide #54
XX Sequence 9 AA;
XX Query Match 88.2%; Score 30; DB 4; Length 9;
XX Best Local Similarity 75.0%; Pred. No. 1.4e+06;
XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 GVXTXIRP 9
Db 2 GVITSIRP 9
XX 2 GVXTXIRP 9
XX 2 GVITSIRP 9
XX 04-DEC-2001 (first entry)
XX Antiangiogenic peptide #54 useful for inhibiting angiogenesis.
XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;
XX arthritis; skin disease; ocular disease; diabetic retinopathy;
XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;
XX cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.
XX Synthetic.
XX Key Location/Qualifiers
FH

```

XX Cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;  
 KW dermatological; immunosuppressive; cardiac; vulvar; antitumor;  
 KW antiatherosclerotic; angiogenesis inhibitor; cancer; arthritis;  
 KW psoriasis; angiogenesis; eye; infection; surgical intervention;  
 KW macular degeneration; diabetic retinopathy; autoimmune disease;  
 KW ocular disease; skin disease; blood vessel disease; telangiectasia;  
 KW Osler Weber Syndrome; myocardial angiogenesis; haemophilic joint;  
 KW plaque neovascularisation; angiofibroma; wound granulation; scleroderma;  
 KW atherosclerosis; intestinal adhesion; Crohn's disease; hypertrophic scar;  
 KW birth control agent; cat scratch disease; ulcer; angiogenic.  
 XX Synthetic.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 XX Modified-site 1  
 FT /label= MeGly  
 FT /note= "This sarcosine residue is modified by N-Ac"  
 FT  
 FT Misc-difference 4  
 FT /note= "This is a D-form residue"  
 FT Modified-site 5  
 FT /note= "Residue is modified to become allophr"  
 FT Modified-site 9  
 FT /note= "Residue is modified by NHCH2CH3"  
 FT  
 XX W0200283065-A2.  
 PN  
 XX  
 XX 24-OCT-2002.  
 PD  
 XX  
 XX 10-APR-2002; 2002WO-US011027.  
 PF  
 XX  
 XX 11-APR-2001; 2001US-00832733.  
 PR  
 XX (ABBO ) ABBOTT LAB.  
 FA  
 XX Henkin J, Haviv F, Bradley MF, Douglas KM, Schneider AJ;  
 PI  
 XX WPI; 2003-111805/10.  
 DR  
 XX New peptides are angiogenesis inhibitors used for treating e.g. cancer,  
 FT arthritis and psoriasis.  
 FT  
 XX Claim 12; Page 31; 33pp; English.  
 PS  
 XX The invention relates to novel peptides for use as angiogenesis  
 CC inhibitors. Used as angiogenesis inhibitors used for treating cancer,  
 CC arthritis, psoriasis, angiogenesis of the eye associated with infection  
 CC or surgical intervention, macular degeneration and diabetic retinopathy.  
 CC The novel peptides are also used for treating autoimmune diseases, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler Weber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma, wound granulation, diseases  
 CC characterised by excessive or abnormal stimulation of endothelial cells  
 CC including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars. The novel peptides are also used as  
 CC birth control agents and for treating cat scratch disease and ulcers.  
 CC This sequence represents one of the angiogenic peptides of the invention  
 CC  
 XX Sequence 9 AA;  
 PS  
 XX Query Match 88.2%; Score 30; DB 6; Length 9;  
 CC Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 CC Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 QY 2 GVXTXIRP 9  
 DB 2 GVITSIRP 9  
 DB  
 XX  
 XX RESULT 8  
 XX AAU14987  
 XX ID AAU14987 standard; peptide; 10 AA.  
 XX

AC AAU14987;  
 XX  
 XX 04-DEC-2001 (first entry)  
 DT  
 DE Antiangiogenic peptide #14 useful for inhibiting angiogenesis.  
 XX  
 XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytotatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX Synthetic.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 XX Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT  
 FT Modified-site 4  
 FT /note= "D-form residue, alloisoleucine"  
 FT Modified-site 6  
 FT /note= "N-Methyl serine"  
 FT Modified-site 10  
 FT /note= "D-form residue, C-terminal amide"  
 FT  
 XX W0200138397-A1.  
 PN  
 XX  
 XX 31-MAY-2001.  
 PD  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 PF  
 XX  
 XX 22-NOV-1999; 99US-00447099.  
 PR  
 XX 31-OCT-2000; 2000US-00702649.  
 PR  
 XX (ABBO ) ABBOTT LAB.  
 FA  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 PI  
 XX WPI; 2001-521804/57.  
 DR  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 FT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 FT diabetic retinopathy.  
 FT  
 XX Claim 38; Page 88; 95pp; English.  
 PS  
 XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers, e.g.  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases, e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #14  
 CC  
 XX Sequence 10 AA;  
 PS  
 XX Query Match 88.2%; Score 30; DB 4; Length 10;  
 CC Best Local Similarity 87.5%; Pred. No. 0.83;  
 CC Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 CC

QY 2 GVXTXIRP 9  
 DB 2 GVXTSIRP 9

RESULT 9  
 AAU14993  
 ID AAU14993 standard; peptide; 10 AA.  
 XX  
 AC AAU14993;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX

Antiangiogenic peptide #20 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= Megly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Misc-difference 4  
 FT /note= "D-form residue"  
 FT Modified-site 6  
 FT /note= "N-Methyl serine"  
 FT Modified-site 10  
 FT /note= "D-form residue, C-terminal amide"

XX WO200138397-A1.  
 PN  
 XX 31-MAY-2001.  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 XX  
 PR 22-NOV-1999; 99US-00447099.  
 PR 31-OCT-2000; 2000US-00702649.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 PI WPI; 2001-521804/57.  
 XX  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 PT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 PT diabetic retinopathy.  
 XX  
 XX Claim 38; Page 88; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Webber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodioma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved

CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #20  
 XX  
 SQ Sequence 10 AA;

Query Match 88.2%; Score 30; DB 4; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 0.83; Mismatches 0; Gaps 0;  
 Matches 6; Conservative 0; Indels 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 DB 2 GVXTSIRP 9

RESULT 10  
 ABP57616  
 ID ABP57616 standard; protein; 139 AA.  
 XX  
 AC ABP57616;  
 XX  
 DT 29-APR-2003 (first entry)  
 XX  
 DE S. murayamaensis ATCC 21414 kinamycin protein SEQ ID NO:17.  
 XX  
 KW Streptomyces murayamaensis ATCC 21414; Streptomyces murayamaensis;  
 KW glycosylated kinamycin; kinamycin; type II polyketide; polyketide;  
 KW antibacterial; cytostatic; infection; antibiotic; antitumour;  
 KW electrophilic azo-coupling agent.  
 XX  
 OS Streptomyces murayamaensis.  
 XX WO2003002066-A2.  
 XX  
 XX 09-JAN-2003.  
 XX  
 XX 27-JUN-2002; 2002WO-US020719.  
 XX  
 XX 27-JUN-2001; 2001US-0301401P.  
 XX  
 XX (DIVE-) DIVERSA CORP.  
 XX  
 XX Short JM, Paradkar A, Varoglu M, Mathur EJ;  
 PI WPI; 2003-210195/20.  
 XX  
 XX N-PSDB; ABZ71139.  
 XX  
 XX New isolated polyketide used e.g. as antibiotic and antitumor agents  
 PT comprises kinamycin molecule comprising at least one saccharide group.  
 XX  
 XX Claim 75; Page 91-92; 119pp; English.

XX The present invention describes a polyketide comprising a kinamycin  
 CC molecule comprising at least one saccharide group. ABZ71132 to ABZ71163  
 CC encode glycosylated kinamycins ABP57609 to ABP57640 isolated from  
 CC Streptomyces murayamaensis ATCC 21414. Kinamycins are a class of type II  
 CC polyketides. The kinamycins have antibacterial and cytostatic activities.  
 CC They can be used for treating infections as antibiotics and as antitumour  
 CC agents, and as electrophilic azo-coupling agents in vitro or in vivo  
 XX  
 SQ Sequence 139 AA;

Query Match 88.2%; Score 30; DB 6; Length 139;  
 Best Local Similarity 75.0%; Pred. No. 15;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 DB 28 GVLTIRP 35

RESULT 11  
 ABU79133  
 ID ABU79133 standard; protein; 2796 AA.



XX ABU79133;  
AC 18-JUN-2003 (first entry)  
DT Mycobacteria mycolic acid protein.  
DE  
XX Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
KW gene therapy; mammalian cell receptor; cytostatic;  
KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
KW tumouricidal immunocyte; antitumour.  
XX  
OS Mycobacteria sp.  
XX  
XX US2002177551-A1.  
XX  
XX 28-NOV-2002.  
XX  
XX 30-MAY-2001; 2001US-00870759.  
XX  
XX 31-MAY-2000; 2000US-0208128P.  
XX  
XX (TERM/) TERMAN D S.  
XX  
XX Terman DS;  
XX  
XX WPI: 2003-361759/34.  
XX N-PSDB; ACA64735.  
XX  
XX A mammalian cell receptor, useful in the treatment of cancer by binding  
XX to tumor associated lipids where the binding induces energy or apoptosis  
XX in T cells and antigen presenting cells.  
XX  
XX Example 2; Page: 167pp; English.  
XX  
XX The invention relates to a mammalian cell receptor, useful in the  
XX treatment of cancer, which binds to tumour associated lipids and induces  
XX energy or apoptosis in the T cells and antigen presenting cells (APCs).  
XX Also included are a mammalian cell useful in the treatment of cancer  
XX where the receptor which binds tumour associated lipids and induces  
XX cellular inactivation or death is deleted or functionally deactivated,  
XX producing (M1) a tumouricidal immunocyte population in vivo in a mammal  
XX (by allowing tumour associated lipids to contact immunocytes in which  
XX receptors for immunosuppressive fatty acids, ceramides, glycolipids,  
XX sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,  
XX sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
XX deleted), a construct useful in the treatment of cancer comprising a  
XX superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell  
XX useful in the treatment of cancer (where an adaptor protein which  
XX inhibits T cell activation by tumour associated antigens is deleted or  
XX functionally deactivated), a composition useful in the treatment of  
XX cancer (comprising a lipid raft conjugated to a superantigen), producing  
XX (M2) a tumouricidal immunocyte population ex vivo in a mammal (by  
XX allowing tumour associated lipids to contact immunocytes, in which  
XX receptors for the lipids are inactivated or deleted to produce a  
XX tumouricidal immunocyte population, and administering the tumouricidal  
XX activated immunocytes to the host), producing (M3) a tumouricidal APC  
XX population ex vivo in a mammal (by allowing a tumour associated lipid to  
XX contact APCs, in which receptors for the tumour associated lipids are  
XX inactivated or deleted to produce a tumouricidally activated population,  
XX and administering APCs to the host), producing a tumouricidal T cell  
XX population ex vivo in a mammal (by allowing a tumour associated lipids to  
XX contact T cells, in which adaptor proteins, which inhibit T cell  
XX activation by tumour associated antigens, are deleted or functionally  
XX deactivated to produce a tumouricidal population of T cells, and  
XX administering the tumouricidally activated T cells to the host, or  
XX allowing a superantigen-lipid raft to contact T cells ex vivo, and  
XX administering the tumouricidally activated T cells to the host), treating  
XX (M5) cancer in a mammal (by administering a lipid binding molecule which  
XX binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
XX a tumouricidal T cell population in vivo in a mammal (by allowing a  
XX tumour associated antigen to contact immunocytes in which adaptor  
XX proteins which inhibit T cell activation by tumour associated antigens

CC are deleted or functionally deactivated) and producing (M7) a  
CC tumouricidal T cell population ex vivo in a mammal comprising allowing a  
CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
CC receptors, methods and compositions are useful for treating cancers and  
CC tumours. Bacterial superantigens are co-administered or administered as  
CC fusion constructs with anti-tumour proteins or motifs. The present  
CC sequence represents an anti-tumour protein which is co-administered with  
CC or incorporated into a fusion construct with a superantigen. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from the US patent  
CC office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"  
XX  
SQ Sequence 2796 AA;  
  
Query Match 88.2%; Score 30; DB 6; Length 2796;  
Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
|||  
Db 2704 GVHTSIRP 2711  
  
RESULT 12  
ADD31274  
ID ADD31274 standard; peptide; 8 AA.  
XX AC ADD31274;  
XX  
XX 15-JAN-2004 (first entry)  
XX DE Angiogenesis inhibiting peptide #49.  
XX  
XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1 /label= OTHER  
FT Misc-difference 3 /note= "OTHER= N-acetyl"  
FT Modified-site 5 /note= "D-form residue"  
FT Modified-site 5 /label= OTHER  
FT Modified-site 8 /note= "OTHER= Nme Norvalyl (Nva)"  
FT /note= "NHCH2CH3"  
XX US2003109455-A1.  
XX 12-JUN-2003.  
XX 30-OCT-2002; 2002US-00283550.  
XX 31-OCT-2001; 2001US-0335017P.  
XX (HAVI/) HAVIV F.  
XX (BRAD/) BRADLEY M F.  
XX Haviv F, Bradley MF;  
XX WPI: 2003-843:01/78.  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX Claim 11; Page 25; 26pp; English.  
PS

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 SQ Sequence 8 AA;  
 Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 2 GVXTXIRP 9  
 Db 1 GVITXIRP 8  
 RESULT 13  
 ADD31265  
 ID ADD31265 standard; peptide; 8 AA.  
 XX  
 AC ADD31265;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Angiogenesis inhibiting peptide #40.  
 XX  
 KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT  
 FT Modified-site 3  
 FT /label= OTHER  
 FT /note= "OTHER= D-form residue, alle (not defined)"  
 FT  
 FT Modified-site 5  
 FT /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT  
 FT Modified-site 8  
 FT /note= "NHCH2CH3"  
 FT  
 FT US2003109455-A1.  
 XX  
 PN 12-JUN-2003.  
 XX  
 PD 30-OCT-2002; 2002US-00283550.  
 XX  
 PF 31-OCT-2001; 2001US-0335017P.  
 XX  
 PR (HAVI/) HAVIV F.  
 XX (BRAD/) BRADLEY M F.  
 PA  
 PI HAVIV F, Bradley MF;  
 XX  
 FI WPI; 2003-843101/78.  
 XX  
 DR New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 XX angiogenesis, treating cancer in mammal, or for prevention of other  
 XX diseases such as autoimmune diseases.  
 PT  
 PT Claim 11; Page 24; 26pp; English.  
 XX  
 PS The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and

CC The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 SQ Sequence 8 AA;  
 Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 GVXTXIRP 9  
 Db 1 GVXTXIRP 8  
 RESULT 14  
 ADD31264  
 ID ADD31264 standard; peptide; 8 AA.  
 XX  
 AC ADD31264;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Angiogenesis inhibiting peptide #39.  
 XX  
 KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT  
 FT Misc-difference 3  
 FT /note= "D-form residue"  
 FT  
 FT Modified-site 5  
 FT /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT  
 FT Modified-site 8  
 FT /note= "NHCH2CH3"  
 FT  
 FT US2003109455-A1.  
 XX  
 PN 12-JUN-2003.  
 XX  
 PD 30-OCT-2002; 2002US-00283550.  
 XX  
 PF 31-OCT-2001; 2001US-0335017P.  
 XX  
 PR (HAVI/) HAVIV F.  
 XX (BRAD/) BRADLEY M F.  
 PA  
 PI HAVIV F, Bradley MF;  
 XX  
 FI WPI; 2003-843101/78.  
 XX  
 DR New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 XX angiogenesis, treating cancer in mammal, or for prevention of other  
 XX diseases such as autoimmune diseases.  
 PT  
 PT Claim 11; Page 24; 26pp; English.  
 XX  
 PS The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and

CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 SQ Sequence 8 AA;

Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 |||||  
 Db 1 GVITXIRP 8

RESULT 15  
 ADD31266  
 ID ADD31266 standard; peptide; 8 AA.  
 XX  
 AC ADD31266;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Angiogenesis inhibiting peptide #41.  
 XX  
 KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.

Key	Location/Qualifiers
FT Modified-site 1	/label= OTHER
FT	/note= "OTHER= N-acetyl"
FT Misc-difference 3	
FT	/note= "D-form residue"
FT Modified-site 4	
FT	/label= OTHER=
FT	/note= "OTHER= alloThr (not defined)"
FT Modified-site 5	
FT	/label= OTHER
FT	/note= "OTHER= Norvalyl (Nva)"
FT Modified-site 8	
FT	/note= "NECH2CH3"

US2003109455-A1.  
 12-JUN-2003.  
 30-OCT-2002; 2002US-00283550.  
 31-OCT-2001; 2001US-0335017P.  
 (HAVI/) HAVIV F.  
 (BRAD/) BRADLEY M F.  
 Haviv F, Bradley MF;  
 WPI; 2003-843101/78.  
 New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 angiogenesis, treating cancer in mammal, or for prevention of other  
 diseases such as autoimmune diseases.  
 Claim 11; Page 24; 26pp; English.  
 The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)

CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 SQ Sequence 8 AA;

Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 |||||  
 Db 1 GVITXIRP 8

Search completed: April 1, 2004, 17:35:24  
 Job time : 43.1053 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:30:38 ; Search time 12.7895 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-4

Perfect score: 34

Sequence: 1 XGVXIXRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	240	2	US-08-114-555A-6
2	29	85.3	240	2	US-08-114-555A-8
3	29	85.3	240	3	US-08-559-397A-12
4	29	85.3	240	3	US-08-559-397A-14
5	29	85.3	748	4	US-09-252-991A-21696
6	28	82.4	271	4	US-09-252-991A-26235
7	27	79.4	1208	4	US-09-540-236-3604
8	26	76.5	216	4	US-09-252-991A-23468
9	26	76.5	479	4	US-09-252-991A-23144
10	26	76.5	728	4	US-09-252-991A-31891
11	25	73.5	208	4	US-09-252-991A-23734
12	25	73.5	293	3	US-09-307-621-4
13	25	73.5	307	4	US-09-252-991A-19676
14	25	73.5	449	1	US-08-476-008-5
15	25	73.5	449	1	US-08-476-008-7
16	25	73.5	449	1	US-08-306-063-5
17	25	73.5	449	1	US-08-306-063-7
18	25	73.5	449	1	US-08-833-485-5
19	25	73.5	449	1	US-08-833-485-7
20	25	73.5	449	3	US-09-137-440-5
21	25	73.5	449	3	US-09-137-440-7
22	25	73.5	449	5	PCT-US91-06148A-5
23	25	73.5	449	5	PCT-US91-06148A-7
24	25	73.5	596	4	US-09-252-991A-18875
25	25	73.5	836	4	US-09-252-991A-26065
26	25	73.5	929	4	US-09-252-991A-19203
27	25	73.5	947	4	US-09-719-190-2

28	25	73.5	1048	4	US-09-921-099A-11	Sequence 11, Appl
29	24	70.6	40	2	US-08-919-724-5	Sequence 5, Appl1
30	24	70.6	60	4	US-09-107-532A-6833	Sequence 6833, Ap
31	24	70.6	62	4	US-09-621-976-4394	Sequence 4394, Ap
32	24	70.6	76	4	US-09-252-991A-25888	Sequence 25888, A
33	24	70.6	95	3	US-09-024-023-4	Sequence 4, Appl1
34	24	70.6	95	4	US-09-531-111-4	Sequence 4, Appl1
35	24	70.6	100	4	US-09-543-681A-5475	Sequence 5475, Ap
36	24	70.6	146	4	US-09-489-039A-13846	Sequence 13846, A
37	24	70.6	161	4	US-09-230-196-8	Sequence 8, Appl1
38	24	70.6	176	4	US-09-252-991A-30148	Sequence 30148, A
39	24	70.6	220	4	US-09-011-151-13	Sequence 13, Appl1
40	24	70.6	229	3	US-09-024-023-2	Sequence 2, Appl1
41	24	70.6	229	4	US-09-531-111-2	Sequence 2, Appl1
42	24	70.6	252	4	US-09-328-352-8187	Sequence 8187, Ap
43	24	70.6	269	2	US-08-727-311-3	Sequence 3, Appl1
44	24	70.6	308	4	US-09-252-991A-28762	Sequence 28762, A
45	24	70.6	323	4	US-09-489-039A-7409	Sequence 7409, Ap

#### ALIGNMENTS

RESULT 1  
US-08-114-555A-6  
; Sequence 6, Application US/08114555A  
; Patent No. 5854392  
; GENERAL INFORMATION:  
; APPLICANT: Manly, Susan P.  
; APPLICANT: Kozlowski, Michael R.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100  
; TITLE OF INVENTION: RECEPTOR (C100-R)  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/114,555A  
; FILING DATE: 30-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/POCKET NUMBER: 6013-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-114-555A-6

Query Match 85.3%; Score 29; DB 2; Length 240;  
Best Local Similarity 62.5%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVTXIXRP 9

```

Db      3  GVTTSLRP 10

RESULT 2
US-08-114-555A-8
; Sequence 8, Application US/08114555A
; Patent No. 5854392
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF
; TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/114,555A
; FILING DATE: 30-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-3090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-114-555A-8

Query Match      85.3%; Score 29; DB 2; Length 240;
Best Local Similarity 82.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels

Qy      2  GVXTXIRP 9
      |||:||
Db      3  GVTTSLRP 10

RESULT 3
US-08-559-397A-12
; Sequence 12, Application US/08559397A
; Patent No. 6083713
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF
; TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA

```

;  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-559-397A-14

Query Match 85.3%; Score 29; DB 3; Length 240;  
Best Local Similarity 62.5%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
Db 3 GVTSLRP 10

#### RESULT 5

US-09-252-991A-21696  
; Sequence 21696, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21696

; LENGTH: 748

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21696

Query Match 85.3%; Score 29; DB 4; Length 748;  
Best Local Similarity 75.0%; Pred. No. 39;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
Db 740 GVTXIRP 747

#### RESULT 6

US-09-252-991A-26235  
; Sequence 26235, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26235

; LENGTH: 271

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26235

Query Match 82.4%; Score 28; DB 4; Length 271;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
Db 85 GVTSLRP 92

#### RESULT 7

US-09-540-236-3604

; Sequence 3604, Application US/09540236

; Patent No. 6673910

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236

; CURRENT FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 3840

; SEQ ID NO 3604

; LENGTH: 1208

; TYPE: PRT

; ORGANISM: M.catarrhalis

US-09-540-236-3604

Query Match 79.4%; Score 27; DB 4; Length 1208;  
Best Local Similarity 50.0%; Pred. No. 28+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
Db 1113 GIITQVRP 1120

#### RESULT 8

US-09-252-991A-23468

; Sequence 23468, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23468

; LENGTH: 216

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23468

Query Match 76.5%; Score 26; DB 4; Length 216;  
Best Local Similarity 50.0%; Pred. No. 60;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
Db 69 GVTALRP 76

#### RESULT 9

US-09-252-991A-23144

; Sequence 23144, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

```

; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23144
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23144

Query Match          76.5%; Score 26; DB 4; Length 479;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      219 GVSSAVRP 226

RESULT 10
US-09-252-991A-31891
; Sequence 31891, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31891
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31891

Query Match          76.5%; Score 26; DB 4; Length 728;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      474 GVSSAVRP 481

RESULT 11
US-09-252-991A-23734
; Sequence 23734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23734
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23734

```

```

Query Match          73.5%; Score 25; DB 4; Length 208;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      5 GVSSAVRP 12

RESULT 12
US-09-307-621-4
; Sequence 4, Application US/09307621
; Patent No. 6066487
; GENERAL INFORMATION:
; APPLICANT: Xiao, Jian-ping
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BarFI Restriction
; TITLE OF INVENTION: Endonuclease In E. coli
; FILE REFERENCE: Bs-PfI
; CURRENT APPLICATION NUMBER: US/09/307,621
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-307-621-4

Query Match          73.5%; Score 25; DB 3; Length 293;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      196 GLKTSIRP 203

RESULT 13
US-09-252-991A-19676
; Sequence 19676, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19676
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19676

Query Match          73.5%; Score 25; DB 4; Length 307;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      23 GVGTTARP 30

RESULT 14
US-08-476-008-5
; Sequence 5, Application US/08476008

```

```
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-008-5

Query Match 73.5%; Score 25; DB 1; Length 449;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 189 GVTTVIEP 196

RESULT 15
US-08-476-008-7
; Sequence 7, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
```

```
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,008
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 449 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-008-7

Query Match 73.5%; Score 25; DB 1; Length 449;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 189 GVTTVIEP 196

Search completed: April 1, 2004, 17:42:12
Job time : 12.7895 secs
```



Sequence 5372, Ap  
Sequence 6787, Ap  
Sequence 126, App  
Sequence 2, Appli  
Sequence 4322, Ap  
Sequence 6, Appli  
Sequence 9, Appli  
Sequence 16984, A  
Sequence 280, App  
Sequence 292, App  
Sequence 9282, Ap  
Sequence 9524, Ap  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 2289, Ap  
Sequence 37683, A  
Sequence 57365, A  
Sequence 16292, A  
Sequence 83, Appl  
Sequence 1552, A  
Sequence 15925, A  
Sequence 17112, A  
Sequence 17562, A  
Sequence 1902, Ap  
Sequence 275517,  
Sequence 227151,  
Sequence 510, App  
Sequence 510, App  
Sequence 510, App  
Sequence 1422, Ap

Sequence 5372, Ap  
Sequence 6787, Ap  
Sequence 126, App  
Sequence 2, Appli  
Sequence 4322, Ap  
Sequence 6, Appli  
Sequence 9, Appli  
Sequence 16984, A  
Sequence 280, App  
Sequence 292, App  
Sequence 9282, Ap  
Sequence 9524, Ap  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 2289, Ap  
Sequence 37683, A  
Sequence 57365, A  
Sequence 16292, A  
Sequence 83, Appl  
Sequence 1552, A  
Sequence 15925, A  
Sequence 17112, A  
Sequence 17562, A  
Sequence 1902, Ap  
Sequence 275517,  
Sequence 227151,  
Sequence 510, App  
Sequence 510, App  
Sequence 510, App  
Sequence 1422, Ap

ALIGNMENTS

RESULT 1  
US-10-424-599-188130  
; Sequence 188130, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188130  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140896C.1.pep  
US-10-424-599-188130

Query Match 88.2%; Score 30; DB 12; Length 89;  
Best Local Similarity 75.0%; Pred. No. 8.2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXIXIRP 9  
DB 78 GVITSIRP 85  
  
RESULT 2  
US-10-187-267A-17  
; Sequence 17, Application US/10187267A  
; Publication No. US20030124679A1  
; GENERAL INFORMATION:  
; APPLICANT: Short, Jay M.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
  
OM protein - protein search, using sw model  
  
Run on: April 1, 2004, 17:35:28 ; Search time 30.6316 Seconds  
(without alignments)  
77.074 Million cell updates/sec  
  
Title: US-09-833-196-4  
Perfect score: 34  
Sequence: 1 XGVXIXIRP 9  
  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
  
Searched: 1069545 seqs, 262320428 residues  
  
Total number of hits satisfying chosen parameters: 1069545  
  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
  
Database : Published Applications AA: \*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep: \*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep: \*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep: \*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep: \*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep: \*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep: \*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep: \*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep: \*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep: \*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep: \*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep: \*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep: \*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep: \*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep: \*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep: \*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep: \*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep: \*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep: \*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	88.2	89	12	US-10-424-599-188130
2	30	88.2	139	14	US-10-187-267A-17
3	30	88.2	2796	9	US-09-870-759-114
4	30	88.2	2796	10	US-09-751-708A-114
5	29	85.3	204	14	US-10-029-386-32508
6	29	85.3	312	14	US-10-156-761-11045
7	29	85.3	314	9	US-09-738-626-6809
8	29	85.3	386	12	US-10-424-599-269274
9	29	85.3	5081	12	US-10-276-774-1850
10	28	82.4	90	9	US-09-864-761-45189
11	28	82.4	147	9	US-09-864-761-45189
12	28	82.4	236	14	US-10-029-386-32508
13	28	82.4	309	15	US-10-264-237-2274
14	28	82.4	332	12	US-10-425-114-46836
15	28	82.4	376	14	US-10-156-761-8348

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```
; ORGANISM: Mycobacterium bovis
US-09-751-708A-114

Query Match      88.2%; Score 30; DB 10; Length 2796;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      ||| |||
Db      2704 GVHTSIRP 2711

RESULT 5
US-10-029-386-32508
; Sequence 32508, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32508
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011469.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: P21817, EVALUE 1.00e-104
US-10-029-386-32508

Query Match      85.3%; Score 29; DB 14; Length 204;
Best Local Similarity 62.5%; Pred. No. 35;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      ||| |||
Db      121 GVHTSLRP 128

RESULT 6
US-10-156-761-11045
; Sequence 11045, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11045
; TYPE: PRT
```

```
; APPLICANT: Paradkar, Ashish
; APPLICANT: Varoglu, Mustafa
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF
; TITLE OF INVENTION: MAKING AND USING THEM
; FILE REFERENCE: 09010-280001
; CURRENT APPLICATION NUMBER: US/10/187,267A
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: US 60/301,401
; PRIOR FILING DATE: 2001-06-27
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Streptomyces murayamaensis ATCC 21414
US-10-187-267A-17

Query Match      88.2%; Score 30; DB 14; Length 139;
Best Local Similarity 75.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      ||| |||
Db      28 GVLTIIRP 35

RESULT 3
US-09-870-759-114
; Sequence 114, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 114
; LENGTH: 2796
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-09-870-759-114

Query Match      88.2%; Score 30; DB 9; Length 2796;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      ||| |||
Db      2704 GVHTSIRP 2711

RESULT 4
US-09-751-708A-114
; Sequence 114, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 114
; LENGTH: 2796
; TYPE: PRT
```

; LENGTH: 312  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-11045  
  
Query Match 85.3%; Score 29; DB 14; Length 312;  
Best Local Similarity 62.5%; Pred. No. 54;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
Db 71 GVRTTVP 78  
  
RESULT 7  
US-09-738-626-6809  
; Sequence 6809, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent in ver. 3.0  
; SEQ ID NO 6809  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6809  
  
Query Match 85.3%; Score 29; DB 9; Length 314;  
Best Local Similarity 62.5%; Pred. No. 55;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
Db 46 GITTFRP 53  
  
RESULT 8  
US-10-424-599-269274  
; Sequence 269274, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 269274  
; LENGTH: 386

; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(386)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_85175C.1.pep  
US-10-424-599-269274  
  
Query Match 85.3%; Score 29; DB 12; Length 386;  
Best Local Similarity 62.5%; Pred. No. 68;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
Db 185 GVATVRP 192  
  
RESULT 9  
US-10-276-774-1850  
; Sequence 1850, Application US/10276774  
; Publication No. US20040053245A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; APPLICANT: Tang, Y, Tom et al  
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 21272-030  
; CURRENT APPLICATION NUMBER: US/10/276,774  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: 09/560,875  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 09/496,914  
; PRIOR FILING DATE: 2000-02-03  
; NUMBER OF SEQ ID NOS: 2700  
; SOFTWARE: Custom  
; SEQ ID NO 1850  
; LENGTH: 5081  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-276-774-1850  
  
Query Match 85.3%; Score 29; DB 12; Length 5081;  
Best Local Similarity 62.5%; Pred. No. 1e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
Db 1809 GVTTSLRP 1816  
  
RESULT 10  
US-09-864-761-45189  
; Sequence 45189, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04

;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 45189  
;; LENGTH: 90  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL049870.3  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.46  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.54  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
;; OTHER INFORMATION: SWISSPROT HIT: Q04652, EVALUE 2.00e-11  
;; OTHER INFORMATION: EST\_HUMAN HIT: N44974.1, EVALUE 7.00e-33  
US-09-864-761-45189

Query Match 82.4%; Score 28; DB 9; Length 90;  
Best Local Similarity 50.0%; Pred. No. 26;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
Db 16 GIATNVRP 23

RESULT 11  
US-09-864-761-46558  
; Sequence 46558, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03

;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 46558  
;; LENGTH: 147  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL049870.1  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6  
;; OTHER INFORMATION: EST\_HUMAN HIT: N44974.1, EVALUE 3.00e-55  
;; OTHER INFORMATION: SWISSPROT HIT: Q04652, EVALUE 1.00e-21  
US-09-864-761-46558

Query Match 82.4%; Score 28; DB 9; Length 147;  
Best Local Similarity 50.0%; Pred. No. 44;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
Db 38 GIATNVRP 45

RESULT 12  
US-10-029-386-32561  
; Sequence 32561, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32561  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Homo sapiens

FEATURE:  
; OTHER INFORMATION: MAP TO AF099810.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
; OTHER INFORMATION: SWISSPROT HIT: P31696, EVALUE 2.00e-17  
US-10-029-386-32561

Query Match 82.4%; Score 28; DB 14; Length 236;  
Best Local Similarity 75.0%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
Db 151 GVPTDIRP 158

RESULT 13

US-10-264-237-2274  
; Sequence 2274, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: Patent In Ver. 3.1  
; SEQ ID NO 2274  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-237-2274

Query Match 82.4%; Score 28; DB 15; Length 309;  
Best Local Similarity 50.0%; Pred. No. 96;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
Db 76 GIATNVRP 83

RESULT 14

US-10-425-114-46836  
; Sequence 46836, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53)1313B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46836  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLM017030E08\_FLI.pep  
US-10-425-114-46836

Query Match 82.4%; Score 28; DB 12; Length 332;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
Db 307 GILTSVRP 314

RESULT 15

US-10-156-761-8348  
; Sequence 8348, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8348  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8348

Query Match 82.4%; Score 28; DB 14; Length 376;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
Db 181 GVTTEVRP 188

Search completed: April 1, 2004, 17:45:39  
Job time : 31.6316 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 9.78947 Seconds  
(without alignments)  
88.434 Million cell updates/sec

Title: US-09-833-196-4

Perfect score: 34

Sequence: 1 XGVXTXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- PIR 78:\*
- 1: Pirl.\*
  - 2: Pirl.\*
  - 3: Pirl.\*
  - 4: Pirl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	377	C82876	oligopeptide trans
2	30	88.2	2796	JC4743	fatty-acid synthas
3	29	85.3	5032	A35041	ryanodine receptor
4	29	85.3	5035	I46646	ryanodine receptor
5	29	85.3	5037	B35041	ryanodine receptor
6	28	82.4	53	H84193	hypothetical prote
7	28	82.4	166	A69820	hypothetical prote
8	28	82.4	215	T44861	probable hydroxyla
9	28	82.4	368	B99563	hypothetical prote
10	28	82.4	401	A69713	septal wall dissol
11	28	82.4	420	T16919	hypothetical prote
12	28	82.4	498	H83793	hypothetical prote
13	27	79.4	141	S09804	hypothetical prote
14	27	79.4	355	D97339	dioxygenase relate
15	27	79.4	496	G87546	acid-CoA ligase, p
16	26	76.5	127	G84653	hypothetical prote
17	26	76.5	277	A39964	photosystem II oxy
18	26	76.5	288	T11966	permease protein o
19	26	76.5	313	F86937	probable pantoate
20	26	76.5	327	I26693	transcription fact
21	26	76.5	340	G70741	hypothetical prote
22	26	76.5	359	T36781	probable gluconate
23	26	76.5	391	T36739	hypothetical prote
24	26	76.5	400	T35334	probable membrane
25	26	76.5	415	I18756	hypothetical prote
26	26	76.5	433	S05654	ND3 intron protein
27	26	76.5	520	T36489	L2 protein - human
28	26	76.5	549	T02790	hypothetical prote
29	26	76.5	597	B71130	probable oligopept

30 26 76.5 868 2 A82722 bifunctional diami  
31 26 76.5 1048 2 C86189 protein t25N20.11  
32 26 76.5 1321 2 A60165 sodium channel pro  
33 26 76.5 4967 2 S72269 ryanodine receptor  
34 26 76.5 4969 2 A37113 ryanodine receptor  
35 25 73.5 77 2 A03343 hypothetical prote  
36 25 73.5 103 2 D71224 hypothetical prote  
37 25 73.5 135 2 T03490 conserved hypothet  
38 25 73.5 287 2 C81066 conserved hypothet  
39 25 73.5 295 1 S26978 succinate dehydrog  
40 25 73.5 296 2 AD3561 citryl-CoA lyase (  
41 25 73.5 300 2 A81804 conserved hypothet  
42 25 73.5 314 2 C75003 na+/Ca+ exchanging  
43 25 73.5 325 2 D71139 probable Na+/Ca2+-  
44 25 73.5 344 2 T29367 cysteine synthase  
45 25 73.5 347 2 A95341 probable hydrolase

## ALIGNMENTS

### RESULT 1

C82876 oligopeptide transport system permease protein UU562 [imported] - Ureaplasma urealyticu  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82876  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
Submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
A:Accession: C82876  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-377 <GLA>  
A:Cross-references: GB:A8002154; GB:AF2222894; NID:G6899557; PIDN:AAF30975.1; GSPDB:GNOC  
A:Experimental source: serovar 3; Biovar 1  
C:Genetics:  
A:Gene: oppB; UU562  
A:Genetic code: SGC3

Query Match 91.2%; Score 31; DB 2; Length 377;  
Best Local Similarity 75.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;  
Matches 6; Conservative 0

QY 2 GVXTXIRP 9  
|||  
DB 146 GVSTSRP 153

### RESULT 2

JC4743 fatty-acid synthase (EC 2.3.1.85) - Mycobacterium bovis  
C:Species: Mycobacterium bovis  
C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 05-May-2000  
C:Accession: JC4743  
R:Fernandes, N.D.; Kolattukudy, P.E.  
Gene 170, 95-99, 1996  
A:Title: Cloning, sequencing and characterization of a fatty acid synthase-encoding gen  
A:Reference number: JC4743; MUID:96200863; PMID:8621098  
A:Accession: JC4743  
A:Molecule type: DNA  
A:Residues: 1-2796 <PER>  
A:Cross-references: GB:U36763; NID:g1036834; PIDN:AA03809.1; PID:g1036835  
A:Note: the source is designated as Mycobacterium tuberculosis var. bovis BCG  
C:Comment: This enzyme catalyzes both de novo synthesis and chain elongation of fatty a  
C:Genetics:  
A:Gene: fas  
C:Superfamily: Mycobacterium tuberculosis fatty-acid synthase  
C:Keywords: acyltransferase; coenzyme A; phosphopantetheine; phosphoprotein  
F:2188-2193/Region: nucleotide binding #status predicted  
F:57/Active site: Ser #status predicted  
F:1693/Binding site: (Ser) (covalent) #status predicted

F:2598/Active site: Cys #status predicted

Query Match 88.2%; Score 30; DB 2; Length 2796;  
Best Local Similarity 75.0%; Pred. No. 52;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 2704 GVHTSIRP 2711

# RESULT 3

A35041  
ryanodine receptor type 1, skeletal muscle - human  
N/Alternate names: calcium release channel protein  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 21-Jan-2000  
C/Accession: A35041; 184622; S66630  
R/Zorato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meissner, G.;  
J. Biol. Chem. 265, 2244-2256, 1990  
A/Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release  
A/Reference number: A35041; MUID:90130482; PMID:2298749  
A/Accession: A35041  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-5032 <ZOR>  
A/Cross-references: GB:J05200; NID:g337721; PIDN:AAA60294.1; PID:g337722  
R/Otsu, K.; Phillips, M.S.; Khanna, V.K.; de Leon, S.; MacLennan, D.H.  
Genomics 13, 835-837, 1992  
A/Title: Refinement of diagnostic assays for a probable causal mutation for porcine and  
A/Reference number: 146644; MUID:92347887; PMID:1639409  
A/Accession: 184622  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 598-722 <RES>  
R/Lynn, S.; Morgan, J.M.; Lamb, H.K.; Meissner, G.; Gillespie, J.I.  
FEBS Lett. 372, 6-12, 1995  
A/Title: Isolation and partial cloning of ryanodine-sensitive Ca(2+) release channel pro  
A/Reference number: S66630; MUID:96032536; PMID:7556644  
A/Accession: S66630  
A/Molecule type: mRNA  
A/Residues: 4690-4968 <LYN>  
A/Experimental source: myometrial smooth muscle  
C/Genetics:  
A/Gene: GDB:RYR1  
A/Cross-references: GDB:120359; OMIM:180901  
A/Map position: 19q13.1-19q13.1  
A/Introns: 642/2  
C/Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo  
C/Keywords: Calcium channel; homotetramer; phosphoprotein; skeletal muscle; transmembran  
F:1788-1984/Domain: transcription initiation factor sigma region 1 homology <SR1>

Query Match 85.3%; Score 29; DB 1; Length 5032;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 1765 GVHTSLRP 1772

# RESULT 4

I46646  
ryanodine receptor, skeletal muscle - pig  
N/Alternate names: calcium release channel protein  
C/Species: Sus scrofa domestica (domestic pig)  
C/Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C/Accession: I46646; S31395; I47133; S26624; A37105; I47212; S18135  
R/Fujii, J.; Otsu, K.; Zorato, F.; de Leon, S.; Khanna, V.K.; Weller, J.E.; O'Brien, P.  
Science 253, 448-451, 1991  
A/Title: Identification of a mutation in porcine ryanodine receptor associated with mali  
A/Reference number: I46645; MUID:91320118; PMID:1862346

A/Accession: I46646

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-5035 <FUJ>

A/Cross-references: GB:M91452; NID:g164647; PIDN:AAA31119.1; PID:g164648

A/Accession: I46645

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-614, 'C', 616-5035 <FU2>

A/Cross-references: GB:M91451; NID:g164645; PIDN:AAA31118.1; PID:g164646

R/Jeab, T.; Brem, G.; Brenig, B.

submitted to the EMBL Data Library, November 1992

A/Description: Genomic organization of porcine skeletal muscle ryanodine receptor gene

A/Reference number: S31395

A/Accession: S31395

A/Molecule type: DNA

A/Residues: 1542-2643 <LEB>

A/Cross-references: EMBL:X69465

R/Jeab, T.; Schmoelz, S.; Brem, G.; Brenig, B.

Genomics 18, 349-354, 1993

A/Title: Genomic organization of the porcine skeletal muscle ryanodine receptor (RYR1)

A/Reference number: A48915; MUID:94117003; PMID:8288238

A/Contents: annotation

R/Harbitz, I.; Kristensen, T.; Bosnes, M.; Kran, S.; Davies, W.

Anim. Genet. 23, 395-402, 1992

A/Title: DNA sequence of the skeletal muscle calcium release channel cDNA and verificat

A/Reference number: I47133; MUID:93036581; PMID:1329581

A/Accession: I47133

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 2-2091, 'A', 2093-3494, 'L', 3496-4163, 'S', 4165-4182, 'R', 4184-4411, 'W', 4413-497

A/Cross-references: EMBL:X62880; NID:g1936; PIDN:CAA44674.1; PID:g1937

R/Harbitz, I.; Kristensen, T.; Kran, S.; Davies, W.

submitted to the EMBL Data Library, August 1992

A/Reference number: S26624

A/Accession: S26624

A/Molecule type: DNA

A/Residues: 482-706 <HAW>

A/Cross-references: EMBL:X68247

R/Harbitz, I.; Chowdhary, B.; Thomsen, P.D.; Davies, W.; Kaufmann, U.; Kran, S.; Gustav

Genomics 9, 243-248, 1990

A/Title: Assignment of the porcine calcium release channel gene, a candidate for the ma

A/Reference number: A37105; MUID:91065640; PMID:2174405

A/Accession: A37105

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 4785-4971, 'R', 4973-5035 <HA2>

A/Cross-references: GB:M32501; NID:g164428; PIDN:AAA31022.1; PID:g164429

R/Ledbetter, N.W.; Freiner, J.K.; Louis, C.F.; Mickelson, J.R.

J. Biol. Chem. 269, 31544-31551, 1994

A/Title: Tissue distribution of ryanodine receptor isoforms and alleles determined by

A/Reference number: A55660; MUID:95081095; PMID:7989322

A/Accession: I47212

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 516-676 <LEB>

A/Cross-references: EMBL:U15965; NID:g562095; PIDN:AAA60467.1; PID:g562096

C/Genetics:

A/Gene: RYR1

A/Introns: 527/1; 559/1; 598/3; 643/2; 1570/3; 1646/2; 1850/3; 1939/3; 2006/3; 2044/1;

A/Note: the list of introns may be incomplete

C/Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo

C/Keywords: calcium channel

Query Match

85.3%; Score 29; DB 1; Length 5035;

Best Local Similarity 62.5%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 1767 GVHTSLRP 1774

RESULT 5

B35041  
ryanodine receptor, skeletal muscle - rabbit  
N:Altrinate names: calcium-release channel protein; junctional channel complex  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 27-Jul-1990 #sequence revision 10-Mar-1994 #text\_change 20-Aug-1999  
C:Accession: S04654; B35041; A36181; S53794; S32504  
R:Takehima, H.; Nishimura, S.; Matsumoto, T.; Ishida, H.; Kangawa, K.; Minamino, N.; Ma  
Nature 339, 439-445, 1989  
A:Title: Primary structure and expression from complementary DNA of skeletal muscle ryan  
A:Reference number: S04654; MUID:83262082; PMID:2725677  
A:Accession: S04654  
A:Molecule type: mRNA  
A:Residues: 1-5037 <TAX>  
A:Cross-references: EMBL:X15750; NID:G1705; PIDN:CAA33762.1; PID:G1710  
A>Note: part of this sequence was confirmed by protein sequencing  
R:Zorato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meissner, G.;  
J. Biol. Chem. 265, 2244-2256, 1990  
A:Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release  
A:Reference number: A35041; MUID:90130482; PMID:2298749  
A:Accession: B35041  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: '1-66','X', '68-2014','D', '2016-3246','E', '3248-3480,3486-4497','LE', '4498-4521','Q', '45  
'4748-4758','N', '4760-5037 <ZOR>  
R:Marka, A.R.; Tempst, P.; Hwang, K.S.; Taubman, M.B.; Inui, M.; Chadwick, C.; Fleischer  
Proc. Natl. Acad. Sci. U.S.A. 86, 8683-8687, 1989  
A:Title: Molecular cloning and characterization of the ryanodine receptor/junctional cha  
A:Reference number: A36181; MUID:90046857; PMID:2813419  
A:Accession: A36181  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA; protein  
A:Residues: 'X', '1223-1235','XX', '1238-1251','1334-1348,1566-1569','X', '1571','X', '1573,1597-1604  
-4679','X', '4681-4689','X', '4691-4693','X', '4695','X', '4697-4700 <MAR>  
A>Note: the proteolytic fragments sequenced here from the junctional channel complex as  
R:Varanyi, M.; Meyer, H.E.  
Biol. Chem. Hoppe-Seyler 376, 45-49, 1995  
A:Title: Sarcoplasmic reticular Ca(2+) release channel is phosphorylated at serine 2843  
A:Reference number: S53794; MUID:95335639; PMID:7612188  
A:Accession: S53794  
A:Molecule type: protein  
A:Residues: 2841-2852 <VAR>  
R:Takehima, H.; Nishimura, S.; Nishi, M.; Ikeda, M.; Sugimoto, T.  
FEBS Lett. 322, 105-110, 1993  
A:Title: A brain-specific transcript from the 3'-terminal region of the skeletal muscle  
A:Reference number: S32504; MUID:93245969; PMID:8097730  
A:Accession: S32504  
A:Molecule type: mRNA  
A:Residues: 4163-5037 <TAX>  
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog  
C:Keywords: calcium channel; glycoprotein; phosphoprotein; skeletal muscle; transmembran  
F:694-704/Region: adenine nucleotide binding  
F:841-954,955-1068/Region: 114-residue repeats  
F:1344-1359,1371-1386/Region: 16-residue repeats  
F:1789-1985/Domain: transcription initiation factor sigma region 1 homology <SRI>  
F:2370-2375/Region: adenine nucleotide binding  
F:2725-2844,2845-2958/Region: 120-residue repeats  
F:4564-4580/Domain: transmembrane #status predicted <TM1>  
F:4641-4664/Domain: transmembrane #status predicted <TM2>  
F:4836-4859/Domain: transmembrane #status predicted <TM3>  
F:4918-4937/Domain: transmembrane #status predicted <TM4>  
F:128,286/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent ki  
F:2843/Binding site: phosphate (Ser) (covalent) #status experimental  
F:4864/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.3%; Score 29; DB 2; Length 5037;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 1766 GVTTSLRP 1773

RESULT 6

H84193  
hypothetical protein Vng0346h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: H84193  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: H84193  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-33 <STO>  
A:Cross-references: GB:AE004437; NID:G10579970; PIDN:AAG18916.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG0346H

Query Match 82.4%; Score 28; DB 2; Length 53;  
Best Local Similarity 62.5%; Pred. No. 2.5;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 34 GVGTGVRP 41

RESULT 7

A59820  
hypothetical protein yhaw - Bacillus subtilis  
C:Species: Bacillus subtilis  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: A59820  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
isch, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.F  
Koetter, P.; Koningsreik, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis.  
A:Title: The complete genome sequence of the Gram-positive bacterium  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: A69820  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-166 <KUN>  
A:Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12823.1; PID:G2633319  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yhaw  
C:Superfamily: Bacillus subtilis hypothetical protein yhaw

Query Match 82.4%; Score 28; DB 2; Length 166;  
Best Local Similarity 62.5%; Pred. No. 8.5;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 108 GILTGIRP 115



akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
A.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: A69713  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-401 <KUN>  
A;Cross-references: GB:Z99117; GB:AL009126; NID:G2634966; PIDN:CAB14495.1; PID:G2634999  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: spoIIP

Query Match 82.4%; Score 28; DB 2; Length 401;  
Best Local Similarity 62.5%; Pred. No. 22;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
DB 38 GVLTSLRP 45

RESULT 11  
Ti6919  
hypothetical protein T21F4.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 18-Feb-2000  
C;Accession: T16919  
R;Minx, P.  
submitted to the EMBL Data Library, April 1996  
A;Description: The sequence of C. elegans cosmid T21F4.  
A;Reference number: Z18605  
A;Accession: T16919  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-420 <MIN>  
A;Cross-references: EMBL:U56959; NID:g1293801; PID:g1293802; PIDN:AAA98709.1; GSPDB:GNO  
A;Experimental source: strain Bristol N2; clone T21F4  
C;Genetics:  
A;Gene: CESP:T21F4.1  
A;Map position: X  
A;Introns: 73/3; 128/3; 185/3; 219/2; 313/3; 386/1

Query Match 82.4%; Score 28; DB 2; Length 420;  
Best Local Similarity 62.5%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
DB 108 GVATAMRP 115

RESULT 12  
H83793  
hypothetical protein BH152 [imported] - Bacillus halodurans (strain C-125)  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C;Accession: H83793  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: H83793  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-498 <STO>  
A;Cross-references: GB:AP001511; GB:BA000004; NID:gi0173727; PIDN:BA004871.1; GSPDB:GNO  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: BH152

Query Match 82.4%; Score 28; DB 2; Length 498;

septal wall dissolution protein spoIIP - Bacillus subtilis  
A;Accession: A69713  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C;Accession: A69713  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Broutillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod

Query Match 82.4%; Score 28; DB 2; Length 368;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
DB 137 GISTALRP 144

RESULT 10  
A69713  
septal wall dissolution protein spoIIP - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C;Accession: A69713  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Broutillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod

Query Match 82.4%; Score 28; DB 2; Length 215;  
Best Local Similarity 62.5%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
DB 185 GVTTLRLP 192

RESULT 9  
B99563  
hypothetical protein MYPV\_4100 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C;Species: Mycoplasma pulmonis  
C;Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C;Accession: B99563  
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A;Reference number: A99512; MUID:21267165; PMID:11353084  
A;Accession: B99563  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-368 <KUR>  
A;Cross-references: GB:AL445566; PID:g14089824; PIDN:CAC13583.1; GSPDB:GNO00153  
A;Experimental source: strain UAB CTIP  
C;Genetics:  
A;Gene: MYPV\_4100  
A;Genetic code: SGC3

Query Match 82.4%; Score 28; DB 2; Length 368;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
||| :||  
DB 137 GISTALRP 144

RESULT 10  
A69713  
septal wall dissolution protein spoIIP - Bacillus subtilis  
C;Species: Bacillus subtilis  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C;Accession: A69713  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Broutillet, S.; Bruchi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mausel  
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod

Best Local Similarity 62.5%; Pred. No. 28;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
DB 107 GVLTGVRP 114

#### RESULT 13

S09804  
Hypothetical protein UL41 - human cytomegalovirus (strain AD169)  
C/Species: human cytomegalovirus, human herpesvirus 5  
A/Note: host Homo sapiens (man)  
C/Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 28-Jul-2000  
C/Accession: S09804  
R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrell, B.G.  
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990  
A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus  
A/Reference number: S09749; MUID:90269039; PMID:2161319  
A/Accession: S09804  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-141 <CHE>  
A/Cross-references: EMBL:X17403; NID:G59591; PIDN:CAA35400.1; PID:e27240; PID:g1813967  
A/Note: this sequence was submitted to the EMBL Data Library, December 1989  
A/Note: this reading frame extends between two stop codons and does not begin with a start  
C/Superfamily: human cytomegalovirus hypothetical protein UL41

Query Match 79.4%; Score 27; DB 2; Length 141;  
Best Local Similarity 50.0%; Pred. No. 13;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 8 GLITLIRP 15

#### RESULT 14

D97339  
dioxigenase related to 2-nitropropane dioxigenase [imported] - Clostridium acetobutylicum  
C/Species: Clostridium acetobutylicum  
C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C/Accession: D97339  
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum  
A/Reference number: A96900; MUID:21359325; PMID:21359325  
A/Accession: D97339  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-355 <KUR>  
A/Cross-references: GB:AE001437; PIDN:AAK81503.1; PID:gl5026676; GSPDB:GN00168  
A/Experimental source: Clostridium acetobutylicum ATCC824  
C/Genetics:  
A/Gene: CAC3580  
C/Superfamily: Campylobacter jejuni hypothetical protein Cj1270c

Query Match 79.4%; Score 27; DB 2; Length 355;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 190 GLVLTAKP 197

#### RESULT 15

G87546  
acid-CoA ligase, probable [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C/Accession: G87546  
R/Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hatt, D.H.; Kolln, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647

A/Accession: G87546  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-496 <STO>  
A/Cross-references: GB:AE005673; NID:gl3423937; PIDN:AAK24371.1; GSPDB:GN00148  
C/Genetics:  
C/Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 79.4%; Score 27; DB 2; Length 496;  
Best Local Similarity 62.5%; Pred. No. 51;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 196 GLITSIRP 203

Search completed: April 1, 2004, 17:40:41  
Job time: 10.7895 secs

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:18 ; Search time 5.52632 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-4

Perfect score: 34

Sequence: 1 XGVXTXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	5035	1 RYR1_PIG	P16960 sus scrofa
2	29	85.3	5037	1 RYR1_RABIT	P11716 oryctolagus
3	29	85.3	5038	1 RYR1_HUMAN	P21817 homo sapien
4	28	82.4	401	1 SP2P_BACSU	P37968 bacillus su
5	28	82.4	1541	1 NX3A_HUMAN	Q3Y4C0 homo sapien
6	27	79.4	141	1 UL41_HCMVA	P16814 human cytom
7	26	76.5	277	1 PSBO_SYNPF	P11472 synchococc
8	26	76.5	313	1 PANC_MYCLE	O69524 mycobacteri
9	26	76.5	340	1 YD60_MYCTU	Q11030 mycobacteri
10	26	76.5	370	1 PIT1_MELGA	Q05749 meleagris g
11	26	76.5	375	1 ADRI_HUMAN	Q96854 homo sapien
12	26	76.5	375	1 ADRI_MOUSE	Q319h1 mus musculu
13	26	76.5	433	1 FMN3_PODAN	P15563 podospora a
14	26	76.5	520	1 VL2_HPV19	P36752 human papil
15	26	76.5	4967	1 RYR2_HUMAN	Q92736 homo sapien
16	26	76.5	4969	1 RYR2_RABIT	P30957 oryctolagus
17	25	73.5	295	1 DHSB_UBTMA	P32420 ustilago ma
18	25	73.5	314	1 O2D3_HUMAN	Q8ngb3 homo sapien
19	25	73.5	449	1 AROA_PSES2	P56952 pseudomonas
20	25	73.5	450	1 AROA_BRUAB	Q9agv2 bruceella ab
21	25	73.5	450	1 AROA_BRUME	Q8yeg1 bruceella me
22	25	73.5	450	1 AROA_BRUSU	Q8g3c4 bruceella su
23	25	73.5	466	1 WR32_ARATH	P59583 arabidopsis
24	25	73.5	511	1 PUR9_SYNY3	P74741 s bifunctio
25	25	73.5	913	1 YCL7_HAFIN	P45114 haemophilus
26	25	73.5	923	1 AGLU_TETPY	Q00906 tetrahymena
27	25	73.5	947	1 Z268_HUMAN	Q14597 homo sapien
28	25	73.5	1017	1 EMIL_MOUSE	Q99x41 mus musculu
29	25	73.5	2335	1 TOR1_SCHPO	O14356 schizosacch
30	25	73.5	2565	1 TRAP_MOUSE	Q80yv3 mus musculu
31	25	73.5	3859	1 TRAP_HUMAN	Q9y4e5 homo sapien
32	24	70.6	134	1 IF2B_PYPAE	Q8zxa6 pyrobaculum
33	24	70.6	143	1 YF46_MYCTU	Q10780 mycobacteri

34 24 70.6 179 1 ATPD\_STAAM  
35 24 70.6 198 1 IFE2\_ARATH  
36 24 70.6 238 1 FLGH\_BUCAP  
37 24 70.6 295 1 PNK\_PSEAE  
38 24 70.6 296 1 PNK\_PSEPK  
39 24 70.6 296 1 PNK\_PSESM  
40 24 70.6 302 1 VG03\_HSV11  
41 24 70.6 318 1 ZNUA\_YERPE  
42 24 70.6 365 1 VHP7\_YEAST  
43 24 70.6 379 1 OXAA\_PROMP  
44 24 70.6 380 1 LEU3\_PHACH  
45 24 70.6 401 1 YQV1\_CABEL

#### ALIGNMENTS

RESULT 1  
RYR1\_PIG  
ID RYR1\_PIG STANDARD; PRT; 5035 AA.  
AC P16960;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RYR1)  
DE (RYR-1) (Skeletal muscle calcium release channel).  
GN RYR1 OR CRC.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]\_TaxID=9823;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Norwegian Landrace; TISSUE=Skeletal muscle;  
RX MEDLINE=93036581; PubMed=1329581;  
RA Harbitz I., Kristensen T., Bosnes M., Kran S., Davies W.;  
RT "DNA sequence of the skeletal muscle calcium release channel cDNA and  
RT verification of the Arg615-->Cys615 mutation, associated with porcine  
RT malignant hyperthermia, in Norwegian landrace pigs.";  
RL Anim. Genet. 23:395-402(1992).  
RN [2]  
RP SEQUENCE OF 1129-2801 FROM N.A.  
RA Brenig B.;  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1129-2643 FROM N.A.  
RC STRAIN=German Landrace; TISSUE=Liver;  
RX MEDLINE=94117003; PubMed=8288238;  
RA Leeb T., Schmolzi S., Brem G., Brenig B.;  
RT "Genomic organization of the porcine skeletal muscle ryanodine  
RT receptor (RYR1) gene coding region 4624 to 7929.";  
RL Genomics 18:349-354(1993).  
RN [4]  
RP SEQUENCE OF 4785-5035 FROM N.A.  
RX MEDLINE=91065640; PubMed=2174405;  
RA Harbitz I., Chowdhary B., Thomsen P.D., Davies W., Kaufman U.,  
RA Kran S., Gustavsson I., Christensen K., Hauge J.G.;  
RT "Assignment of the porcine calcium release channel gene, a candidate  
RT for the malignant hyperthermia locus, to the 6p11-->q21 segment of  
RT chromosome 6.";  
RL Genomics 8:243-248(1990).  
RC -!- FUNCTION: Communication between transverse-tubules and  
CC sarcoplasmic reticulum. Contraction of skeletal muscle is  
CC triggered by release of calcium ions from SR following  
CC depolarization of T-tubules.  
CC -!- SUBUNIT: Homotrimer (Potential).  
CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium  
CC ions, magnesium ions, AIP and calmodulin.  
CC -!- MISCELLANEOUS: The calcium release channel activity resides in the  
CC C-terminal region while the remaining part of the protein  
CC constitutes the 'foot' structure spanning the junctional gap  
CC between the SR and the T-tubule. It is possible that the foot  
CC structure interacts with the cytoplasmic region of the



```

FT CAROHYD 4864 4864 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 2015 E -> D (IN REF. 2).
FT CONFLICT 3481 3485 MISSING (IN REF. 2).
SQ SEQUENCE 5037 AA; 565238 MW; 4ABD87AA26697070 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 5037;
Best Local Similarity 62.5%; Pred. No. 90;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps

QY 2 GVXTXIRP 9
||| :||
Db 1766 GVTTSLRP 1773

RESULT 3
RYR1_HUMAN STANDARD; PRT; 5038 AA.
ID RYR1_HUMAN STANDARD; PRT; 5038 AA.
AC P21817; Q16314; Q16368; Q9NPK1; Q9P104;
DC 01-MAY-1991 (Rel. 18, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RYR1)
DE (RYR-1) (Skeletal muscle calcium release channel).
DS RYR1 OR RYDR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND PARTIAL SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90130482; PubMed=2298749;
RA Zorzato F., Fujii J., Otsu K., Phillips M.S., Green N.M., Lai F.A.,
RA Weisner G., MacLennan D.H.;
RT "Molecular cloning of cDNA encoding human and rabbit forms of the
RT Ca2+ release channel (ryanodine receptor) of skeletal muscle
RT sarcoplasmic reticulum.";
RL J. Biol. Chem. 265:2244-2256(1990).
RN [2]
RP REVISIONS TO 2324; 2840 AND 3380, VARIANT MH ARG-248, AND VARIANTS
RP CYS-471, LEU-1787 AND CYS-2060.
RC TISSUE=Muscle;
RX MEDLINE=92372020; PubMed=1354642;
RA Gillard E.F., Otsu K., Fujii J., Duff C.L., de Leon S., Khanna V.K.,
RA Britt B.A., Worton R.G., McLennan D.H.;
RA "Polymorphisms and deduced amino acid substitutions in the coding
RT sequence of the ryanodine receptor (RYR1) gene in individuals with
RT malignant hyperthermia.";
RL Genomics 13:1247-1354(1992).
RN [3]
RP REVISIONS TO 1365-1368, VARIANT CDT/MH HIS-2435, AND ALTERNATIVE
RP SPLICING.
RC TISSUE=Muscle;
RX MEDLINE=94035117; PubMed=8220422;
RA Zhang Y., Chen H.S., Khanna V.K., de Leon S., Phillips M.S.,
RA Schappert K., Britt B.A., Brownell A.K.W., McLennan D.H.;
RT "A mutation in the human ryanodine receptor gene associated with
RT central core disease.";
RL Nat. Genet. 5:46-50(1993).
RN [4]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=96299657; PubMed=8661021;
RA Phillips M.S., Fujii J., Khanna V.K., de Leon S., Yokobata K.,
RA de Jong P.J., McLennan D.H.;
RT "The structural organization of the human skeletal muscle ryanodine
RT receptor (RYR1) gene.";
RL Genomics 34:24-41(1996).
RN [5]
RP SEQUENCE OF 598-722 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=92347887; PubMed=1639409;
RA Otsu K., Phillips M.S., Khanna V.K., de Leon S., McLennan D.H.;
RT "Refinement of diagnostic assays for a probable causation for
RT
```

RT porcine and human malignant hyperthermia.";  
RL Genomics 13:835-837(1992).  
[6]  
RP SEQUENCE OF 3339-5038 FROM N.A. (ISOFORM 2).  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S.,  
RA Phan K., Velasco N., Do L., Regala W., Terry A., Ganes J.,  
RA Dangnan L., Eler A., Christensen M., Georgescu A., Avila J.,  
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,  
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,  
RA Kobayashi A., Olsen A.S., Carraro A.V.,  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
[7]  
RP SEQUENCE OF 4696-4974 FROM N.A.  
RP TISSUE=Myometrium;  
XC MEDLINE=96032536; PubMed=7556644;  
RX Lynn S., Morgan J.M., Lamb H.K., Meissner G., Gillespie J.I.;  
RA "Isolation and partial cloning of ryanodine-sensitive Ca2+ release  
RT channel protein isoforms from human myometrial smooth muscle.";  
RL FBS Lett. 372:6-12(1995).  
[8]  
RP TISSUE SPECIFICITY.  
RX Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;  
RA "Partial cloning and differential expression of ryanodine  
RT receptor/calcium-release channel genes in human tissues including  
RL the hippocampus and cerebellum.";  
RX Neuroscience 85:205-216(1998).  
[9]  
RP S-NITROSYLATION OF CYS-3635.  
RX MEDLINE=21457268; PubMed=11562475;  
RA Sun J., Xin C., Eu J.P., Stanler J.S., Meissner G.;  
RL "Cysteine-3635 is responsible for skeletal muscle ryanodine receptor  
RT modulation by NO.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:11158-11162(2001).  
[10]  
RP VARIANT MH CYS-614.  
RX MEDLINE=92128959; PubMed=1774074.  
RA Gillard E.F., Otsu K., Fujii J., Khanna V.K., de Leon S.,  
RA Demetzi J., Britt B.A., Duff C.L., Worton R.G., McLeenan D.H.;  
RA "A substitution of cysteine for arginine 614 in the ryanodine  
RT receptor is potentially causative of human malignant hyperthermia.";  
RL Genomics 11:751-755(1991).  
[11]  
RP VARIANTS CCD/MH CYS-163 AND MET-403.  
RX MEDLINE=94035118; PubMed=8220423;  
RA Quane K.A., Healy J.M.S., Keating K.E., Manning B.M., Couch F.J.,  
RA Palmucci L.M., Doriguzzi C., Fagerlund T.H., Berg K., Ording H.,  
RA Bendixen D., Mortier W., Linz U., Muller C.R., McCarthy T.V.;  
RT "Mutations in the ryanodine receptor gene in central core disease and  
RL malignant hyperthermia.";  
RL Nat. Genet. 5:51-55(1993).  
[12]  
RP VARIANT CCD/MH SER-522.  
RX MEDLINE=95130087; PubMed=7829078;  
RA Quane K.A., Keating K.E., Healy J.M.S., Manning B.M.,  
RA Krivosic-Horber R., Krivosic I., Monnier N., Lunardi J.,  
RA McCarthy T.V.;  
RT "Mutation screening of the RYR1 gene in malignant hyperthermia:  
RT detection of a novel Tyr to Ser mutation in a pedigree with  
RT associated central cores.";  
RL Genomics 23:236-239(1994).  
[13]  
RP VARIANT MH ARG-341.  
RX MEDLINE=94282042; PubMed=8012359;  
RA Quane K.A., Keating K.E., Manning B.M., Healy J.M.S., Monseurs K.,  
RA Heffron J.J.A., Lehane M., Heytens L., Krivosic-Horber R., Adnet P.,  
RA Ellis F.R., Monnier N., Lunardi J., McCarthy T.V.;  
RT "Detection of a novel common mutation in the ryanodine receptor gene  
RT in malignant hyperthermia: implications for diagnosis and  
RL heterogeneity studies.";  
RL Hum. Mol. Genet. 3:471-476(1994).  
[14]  
RP VARIANT MH ARG-2434.  
RX MEDLINE=95152512; PubMed=7849712;  
RA Keating K.E., Quane K.A., Manning B.M., Lehane M., Hartung E.,  
RA Censier K., Urwyler A., Klausnitzer M., Muller C.R., Heffron J.J.A.,  
RA McCarthy T.V.;  
RT "Detection of a novel RYR1 mutation in four malignant hyperthermia  
RL pedigrees.";  
RL Hum. Mol. Genet. 3:1855-1858(1994).  
[15]  
RP VARIANT MH ARG-2434.  
RX MEDLINE=95187158; PubMed=7881417;  
RA Phillips M.S., Khanna V.K., de Leon S., Prodis W., Britt B.A.,  
RA McLeenan D.H.;  
RT "The substitution of Arg for Gly2433 in the human skeletal muscle  
RL ryanodine receptor is associated with malignant hyperthermia.";  
RL Hum. Mol. Genet. 3:2181-2186(1994).  
[16]  
RP VARIANT MH CYS-614.  
RX TISSUE=Blood;  
RT MEDLINE=95271229; PubMed=7751854;  
RA Moroni I., Gorano E.F., Comi G.P., Tegazzin V., Prella A., Bordini A.,  
RA Bresolin N., Scarlato G.;  
RT "Ryanodine receptor gene point mutation and malignant hyperthermia  
RL susceptibility.";  
RL J. Neurol. 242:127-133(1995).  
[17]  
RP VARIANT MH ARG-35.  
RX MEDLINE=97219028; PubMed=9066328;  
RA Lynch P.J., Krivosic-Horber R., Reiford H., Monnier N., Quane K.A.,  
RA Adnet P., Haudecoeur G., Krivosic I., McCarthy T.V., Lunardi J.;  
RT "Identification of heterozygous and homozygous individuals with the  
RL novel RYR1 mutation Cys35Arg in a large kindred.";  
RL Anesthesiology 86:620-626(1997).  
[18]  
RP VARIANT MH LEU-614.  
RX MEDLINE=98051290; PubMed=9389851;  
RA Quane K.A., Ording H., Keating K.E., Manning B.M., Heine R.,  
RA Bendixen D., Berg K., Krivosic-Horber R., Lehmann-Horn F.,  
RA Fagerlund T.H., McCarthy T.V.;  
RT "Detection of a novel mutation at amino acid position 614 in the  
RL ryanodine receptor in malignant hyperthermia.";  
RL Br. J. Anaesth. 79:332-337(1997).  
[19]  
RP VARIANT MH TRP-552.  
RX MEDLINE=97284075; PubMed=9138151;  
RA Keating K.E., Gibling L., Lynch P.J., Quane K.A., Lehane M.,  
RA Heffron J.J.A., McCarthy T.V.;  
RT "Detection of a novel mutation in the ryanodine receptor gene in an  
RL Irish malignant hyperthermia pedigree: correlation of the IVCT  
RL response with the affected and unaffected haplotypes.";  
RL J. Med. Genet. 34:291-296(1997).  
[20]  
RP VARIANTS MH CYS-2163, MET-2168 AND MET-2206, AND VARIANT CCD/MH  
HIS-2163.  
RX MEDLINE=98163444; PubMed=9497245;  
RA Manning B.M., Quane K.A., Ording H., Urwyler A., Tegazzin V.,  
RA Lehane M., O'Halloran J., Hartung E., Gibling L.M., Lynch P.J.,  
RA Vaughan P., Censier K., Bendixen D., Comi G.P., Heytens L.,  
RA Monseurs K., Fagerlund T.H., Wolz W., Heffron J.J.A., Mueller C.R.,  
RA McCarthy T.V.;  
RT "Identification of novel mutations in the ryanodine-receptor gene  
RL (RYR1) in malignant hyperthermia: genotype-phenotype correlation.";  
RL Am. J. Hum. Genet. 62:599-609(1998).  
[21]  
RP VARIANTS MH CYS-2459 AND HIS-2459.  
RX MEDLINE=98111378; PubMed=9450902;  
RA Manning B.M., Quane K.A., Lynch P.J., Urwyler A., Tegazzin V.,  
RA Krivosic-Horber R., Censier K., Comi G.P., Adnet P., Wolz W.,  
RA Lunardi J., Muller C.R., McCarthy T.V.;  
RL Query Match 85.3%; Score 29; DB 1; Length 5038;  
Best Local Similarity 62.5%; Pred. No. 90;

```
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 2 GVXTXIRP 9
DB 1766 GVLTSLRP 1773

RESULT 4
SP2P_BACSU STANDARD; PRT; 401 AA.
AC P37958;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Stage II sporulation protein P.
GN SPOIIP OR BSU25530.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Takemaru K.I., Sato T., Kobayashi Y.,
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.,
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the Bacillus subtilis genome containing the skin element and many
RT sporulation genes.";
RL Microbiology 142:3103-3111(1996).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Bourschi C.V., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conner I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Etian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goifeau A., Golligly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Muel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescecun E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter B., Scoffone F., Soldo B.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RA Nature 390:249-256(1997).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D17650; BAA04542.1; -
DR EMBL; D84432; BAA12458.1; -
DR EMBL; Z99117; CAB14495.1; -
DR PIR; A69713; A69713.
DR Subtilist; BG10439; spoIIP.
KW Sporulation; Complete proteome.
SQ SEQUENCE 401 AA; 4548 MW; EB060014088E17A5 CRC64;

Query Match 82.4%; Score 28; DB 1; Length 401;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 38 GVLTSLRP 45

RESULT 5
NX3A_HUMAN STANDARD; PRT; 1541 AA.
AC Q9Y4C0; Q9S378; Q9NS47; Q9P1V3; Q9P1V6; Q9UIE2; Q9UIE3; Q9ULA5;
AC Q9Y486;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neurexin 3-alpha precursor (Neurexin III-alpha).
GN NRXN3 OR KIAA0743
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21945268; PubMed=11944992;
RA Rowen L., Young J., Birditt B., Kaur A., Madan A., Philipps D.L.,
RA Qin S., Minx P., Wilson R.K., Hood L., Graveley B.R.;
RT "Analysis of the human neurexin genes: alternative splicing and the
RT generation of protein diversity.";
RL Genomics 79:587-597(2002).
[2]
RP SEQUENCE OF 253-1541 FROM N.A.
RN TISSUE=Brain;
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XI.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).
CC -!- FUNCTION: Neuronal cell surface protein that may be involved in
CC cell recognition and cell adhesion. May mediate intracellular
CC signaling.
CC -!- SUBUNIT: The laminin G-like domain 2 binds to neurexophilin 1.
CC Specific isoforms bind to alpha-dystroglycan. The cytoplasmic C-
CC terminal region binds to CASK (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative promoter;
CC Comment=A number of isoforms, alpha-type (shown here) and
CC beta-type (AC Q9HDB5), are produced by use of alternative
CC promoters. Beta-type isoforms differ from alpha-type isoforms
CC in their N-terminus;
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms may be produced;
CC Name=1;
CC IsoId=Q9Y4C0-1; Sequence=Displayed;
CC -!- TISSUE SPECIFICITY: Predominantly expressed in brain.
CC -!- SIMILARITY: Contains 6 laminin G-like domains.
```

CC	-!- SIMILARITY: Contains 3 EGF-like domains.
CC	-!- SIMILARITY: Belongs to the neurexin family.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to licenses@isb-sib.ch).
CC	
DR	EMBL; AF09810; AAC68909.1; -
DR	EMBL; AC008045; AAF28465.1; -
DR	EMBL; AC009336; AAF21147.1; -
DR	EMBL; AC012099; AAF15058.1; -
DR	EMBL; AC026888; AAF87841.1; -
DR	EMBL; AC011440; AAF61277.1; -
DR	EMBL; AC008056; AAF09143.1; -
DR	EMBL; AFL23462; AAD13621.1; -
DR	EMBL; AB018286; BAA34463.1; ALT_INIT.
DR	HSSP; Q63373; IC4R.
DR	Genew; HGNC:8010; NRXN3.
DR	MIM; 600567; -
DR	GO; GO:0005887; C:integral to plasma membrane; TAS.
DR	GO; GO:0004872; F:receptor activity; TAS.
DR	GO; GO:0007411; P:axon guidance; TAS.
DR	InterPro; IPR000152; ASX_hydroxyl_S.
DR	InterPro; IPR006209; EGFLike.
DR	InterPro; IPR001791; Laminin_G.
DR	Pfam; PF00008; EGF_3.
DR	Pfam; PF00054; laminin G; 6.
DR	PROSITE; PS00010; ASX_HYDROXYL_1.
DR	PROSITE; PS00022; EGF_1; FALSE_NEG.
DR	PROSITE; PS01186; EGF_2; FALSE_NEG.
DR	PROSITE; PS50026; EGF_3; 3.
DR	PROSITE; PS50025; LAM_G_DOMAIN; 6.
KW	Signal; Transmembrane; Repeat; Cell adhesion; EGF-like domain;
KW	Glycoprotein; Alternative splicing; Alternative promoter usage.
FT	SIGNAL 1 27 BY SIMILARITY.
FT	CHAIN 28 1541 NEUREXIN 3-ALPHA.
FT	DOMAIN 28 1466 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1467 1487 POTENTIAL.
FT	DOMAIN 1488 1541 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 28 202 LAMININ G-LIKE 1.
FT	DOMAIN 198 235 EGF-LIKE 1.
FT	DOMAIN 258 440 LAMININ G-LIKE 2.
FT	DOMAIN 447 639 LAMININ G-LIKE 3.
FT	DOMAIN 643 680 EGF-LIKE 2.
FT	DOMAIN 685 857 LAMININ G-LIKE 4.
FT	DOMAIN 871 1046 LAMININ G-LIKE 5.
FT	DOMAIN 1049 1086 EGF-LIKE 3.
FT	DOMAIN 1090 1260 LAMININ G-LIKE 6.
FT	DOMAIN 1289 1292 POLY-THR.
FT	DOMAIN 1295 1298 POLY-THR.
FT	DOMAIN 1370 1373 POLY-THR.
FT	DOMAIN 1385 1388 POLY-SER.
FT	DOMAIN 1474 1477 POLY-ALA.
FT	CARBOHYD 58 58 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 757 757 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1189 1189 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1257 1257 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1301 1301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 1043 1043 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT 1334 1334 K -> E (IN REF. 1; AAF61277).
FT	CONFLICT 1334 1334 T -> TTGGRVPLFVEDATPATAPATITLPPTFRP
FT	LTIETTKDSMTSEAGLCISDGSDGDDGLVSGY
FT	GSGETFDNSLPPTDEDEFYTFSLVT (IN REF. 1;
FT	AAD13621)
FT	SEQUENCE 1541 AA; 169923 MW; BB67FE2FB6197C95 CRC64;
FT	
Query Match	82.4%; Score 28; DB 1; Length 1541;
Best Local Similarity	75.0%; Pred. No. 52;

```

Matches      6;  Conservative      0;  Mismatches      2;  Indels      0;  Gaps      0;

QY      2 GVXTXIRP 9
DB      151 GVPTDIRP 158

RESULT 6
UL41_HCMVA STANDARD; PRT; 141 AA.
ID _UL41_HCMVA
AC P16814;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Hypothetical protein UL41.
GN UL41.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
RA "Analysis of the protein-coding content of the sequence of human
RT cytomegalovirus strain AD169.";
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: X17403; CAA35400.1; -.
DR PUR: S09804; S09804.
KW Hypothetical protein.
SQ SEQUENCE 141 AA; 16766 MW; 20005377B1BFB712 CRC64;

Query Match      79.4%; Score 27; DB 1; Length 141;
Best Local Similarity 50.0%; Pred.No. 9.4;
Matches      4;  Conservative      2;  Mismatches      2;  Indels      0;  Gaps      0;

QY      2 GVXTXIRP 9
DB      8 GIITLRP 15

RESULT 7
PSBO_SYN7 STANDARD; PRT; 277 AA.
ID _PSBO_SYN7
AC P11472;
DT 01-OCT-1999 (Rel. 12, Created)
DT 01-OCT-1999 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Photosystem II manganese-stabilizing polypeptide precursor (MSP).
GN PSBO OR WXXA.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=1140;
RN [1]
RS SEQUENCE FROM N.A.
RX MEDLINE=89068567; PubMed=3120187;
RA Kuwabara T., Reddy K. J., Sherman L.A.;
RT "Nucleotide sequence of the gene from the cyanobacterium Anacystis
RT nidulans R2 encoding the Mn-stabilizing protein involved in
RT photosystem II water oxidation.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:8230-8234 (1987).
RN [2]

```



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL583917; CAC29738.1; -  
EMBL; AL023093; CAA18788.1; -  
PIR; F86937; F86937.  
Leptoma; ML0230; -  
HAMAP; MF 00158; -; 1.  
InterPro; IPR003721; Pantoate\_ligase.  
Pfam; PF02569; Pantoate\_ligase; 1.  
TIGRFAMs; TIGR00019; panC; 1.  
Pantothenate biosynthesis; Ligase; Complete proteome.  
SEQUENCE 313 AA; 33585 MW; 0209C216EB26A712 CRC64;

Query Match 76.5%; Score 26; DB 1; Length 313;  
Best Local Similarity 50.0%; Pred. No. 37;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
| : | | |  
Db 116 GLRTVTRP 123

RESULT 9  
YD60 MYCTU STANDARD; PRT; 340 AA.  
Q11030;  
Q11030; AC  
01-OCT-1996 (Rel. 34, Created)  
01-OCT-1996 (Rel. 34, Last sequence update)  
10-OCT-2003 (Rel. 42, Last annotation update)  
Hypothetical protein RV1360/MT1405/MB1395 precursor.  
RV1360 OR MT1405 OR MT102B10.24 OR MB1395.  
Mycobacterium tuberculosis, and  
Mycobacterium bovis.  
OS  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
NCBI\_taxID=1773, 1765;  
[1]  
SEQUENCE FROM N.A.  
SPECIES=M.tuberculosis; STRAIN=H37Rv;  
MEDLINE=98295987; PubMed=9634230;  
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,  
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
Hornsbey T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
Ruster S., Seeger K., Skelton S., Squares S., Squares R.,  
Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;  
"Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence";  
Nature 393:537-544(1998).  
[2]  
SEQUENCE FROM N.A.  
SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
MEDLINE=2206494; PubMed=12218036;  
Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
Biswal W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;  
"Whole-genome comparison of Mycobacterium tuberculosis clinical and  
laboratory strains";  
J. Bacteriol. 184:5479-5490(2002).  
[3]  
SEQUENCE FROM N.A.  
SPECIES=M.bovis; STRAIN=AF2122/97;  
MEDLINE=22709107; PubMed=12788972;  
RX

RA Garnier T., Eigemeier K., Camus J.-C., Medina N., Mansoor H.,  
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,  
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,  
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.,  
RA "The complete genome sequence of Mycobacterium bovis",  
RA Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z75555; CAA99965.1; -  
CC EMBL; AE007012; AAK45668.1; -  
CC EMBL; BX248338; CAX94256.1; -  
CC PIR; G70741; G70741.  
CC TIGR; MT1405; -  
CC TubercuList; Kvi360; -  
CC InterPro; IPR002103; Bac\_luciferase.  
CC KW Hypothetical protein; Signal; Complete proteome.  
CC SIGNAL 1 20 POTENTIAL.  
CC CHAIN 21 340 HYPOTHETICAL PROTEIN  
CC FT RVL1360/MT1405/MB1395.  
CC FT  
CC SQ SEQUENCE 340 AA; 37252 MW; F4D35CA7C5AC0767 CRC64;  
Query Match 76.5%; Score 26; DB 1; Length 340;  
Best Local Similarity 62.5%; Pred. No. 40;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
Db 93 GLGTQIRP 100  
RESULT 10  
ID PIT1 MELGA STANDARD; PRT; 370 AA.  
AC Q05749.  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Pituitary-specific positive transcription factor 1 (Pit-1) (Growth  
DE hormone factor 1) (GHF-1).  
GN POU1F1 OR PIT1.  
OS Meleagris gallopavo (Common turkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.  
OX NCBI\_TaxID=9103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pituitary;  
RX MEDLINE=93039571; PubMed=1418622;  
RA Wong E.A., Slijs J.L., el Halawani M.E.;  
RT "Complementary DNA cloning and expression of Pit-1/GHF-1 from the  
RT domestic turkey,";  
RL DNA Cell Biol. 11:651-660(1992).  
RN [2]  
RP REVISIONS, AND ALTERNATIVE SPLICING.  
RA Wong E.A., Sharova L., Kurima K., Weatherly K.L.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Transcription factor that activates growth hormone and  
CC prolactin genes. Specifically binds to the consensus sequence 5'-  
CC TAAAT-3'.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=3;  
CC Name=PIT-1beta\*;  
CC IsoId=Q05749-1; Sequence=Displayed;  
CC Name=PIT-1;  
CC IsoId=Q05749-2; Sequence=VSP\_002318;

CC Name=PIT-1\*;  
CC IsoId=Q05749-3; Sequence=VSP\_002319;  
CC -!- TISSUE SPECIFICITY: Pituitary Gland.  
CC -!- SIMILARITY: Belongs to the POU transcription factor family.  
CC Class=1 subfamily.  
CC -!- SIMILARITY: Contains 1 homeobox domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U18928; AAB04690.1; -  
CC EMBL; U62732; AAB04690.1; JOINED.  
CC EMBL; U18923; AAB04690.1; JOINED.  
CC EMBL; U18924; AAB04690.1; JOINED.  
CC EMBL; U18925; AAB04690.1; JOINED.  
CC EMBL; U18926; AAB04690.1; JOINED.  
CC EMBL; U18927; AAB04690.1; JOINED.  
CC EMBL; U18928; AAB04690.1; JOINED.  
CC EMBL; U62732; AAB04691.1; JOINED.  
CC EMBL; U18923; AAB04691.1; JOINED.  
CC EMBL; U18924; AAB04691.1; JOINED.  
CC EMBL; U18925; AAB04691.1; JOINED.  
CC EMBL; U18926; AAB04691.1; JOINED.  
CC EMBL; U18927; AAB04691.1; JOINED.  
CC EMBL; U18928; AAB04691.1; JOINED.  
CC EMBL; U18923; AAB04692.1; JOINED.  
CC EMBL; U18924; AAB04692.1; JOINED.  
CC EMBL; U18925; AAB04692.1; JOINED.  
CC EMBL; U18926; AAB04692.1; JOINED.  
CC EMBL; U18927; AAB04692.1; JOINED.  
CC EMBL; X59471; CAA49229.1; ALT\_SEQ.  
CC PIR; S26693; S26693.  
CC HSP; P10037; 1AU7.  
CC InterPro; IPR001356; Homeobox.  
CC InterPro; IPR000327; POU domain.  
CC InterPro; IPR007103; POU homeo.  
CC Pfam; PF00046; homeobox; 1.  
CC Pfam; PF00157; pou; 1.  
CC PRINTS; PR00028; POU DOMAIN.  
CC ProDom; PD000010; Homeobox; 1.  
CC ProDom; PD000583; POU domain; 1.  
CC SMART; SM00389; HOX; 1.  
CC SMART; SM00352; POU; 1.  
CC PROSITE; PS00027; HOMEBOX 1; 1.  
CC PROSITE; PS00035; POU 1; 1.  
CC PROSITE; PS00465; POU 2; 1.  
CC PROSITE; PS00711; HOMEBOX 2; 1.  
CC Homeobox; DNA-binding; transcription regulation; Activator;  
CC Nuclear protein; Alternative splicing.  
CC DOMAIN 206 276 POU.  
CC DNA\_BIND 292 351 HOMEBOX.  
CC VARSPPLIC 1 47 MTCQAFASDNEVPLNSDSSPSLPIMHSAECLPVSNHA  
CC TSVST -> MYLESSCVFLP (in isoform PIT-1).  
CC /FTID=VSP\_002318.  
CC Missing (in isoform PIT-1\*).  
CC FT VARSPLIC 48 75 /FTID=VSP\_002319.  
CC FT  
CC SQ SEQUENCE 370 AA; 41191 MW; AB2ABFCBE257C171 CRC64;  
Query Match 76.5%; Score 26; DB 1; Length 370;  
Best Local Similarity 62.5%; Pred. No. 43;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
Db 94 GVMTGIKP 101  
RESULT 11

```

ADRI_HUMAN STANDARD; PRT; 375 AA.
AC Q96A54; Q9Y360;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Adiponectin receptor protein 1 (CGI-45).
GN ADIPOR1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RL Caenorhabditis elegans by comparative proteomics.";
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=lung, and Skin;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenner C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,
RA Rana S.S., Locantello N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Rahvey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 95:16899-16903(2002).
RN [3]
RP FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=22687101; PubMed=12802337;
RA Sugiyama T., Miyagishi M., Hara K., Teunoda M., Murakami K.,
RA Ohteki T., Uchida S., Takekawa S., Waki H., Tsuno N.H., Shibata Y.,
RA Terauchi Y., Froguel P., Toke K., Koyasu S., Taira K., Kitamura T.,
RA Shimizu T., Nagai R., Kadowaki T.;
RT "Cloning of adiponectin receptors that mediate antidiabetic metabolic
RL effects.";
RL Nature 423:762-769(2003).
CC -1- FUNCTION: Receptor for globular and full-length adiponectin
CC (APM1), an essential hormone secreted by adipocytes that acts as
CC an antidiabetic. Probably involved in metabolic pathways that
CC regulate lipid metabolism such as fatty acid oxidation. Mediates
CC increased AMPK, PPARG ligand activity, fatty acid oxidation and
CC glucose uptake by adiponectin. Has some high-affinity receptor for
CC globular adiponectin but low-affinity receptor for full-length
CC adiponectin.
CC -1- SUBUNIT: May form homo and heteromultimers.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Localized to the
CC cell membrane and intracellular organelles.
CC -1- TISSUE SPECIFICITY: Widely expressed. Highly expressed in skeletal
CC muscle. Expressed at intermediate level in brain, heart, spleen,
CC kidney, liver, placenta, lung and peripheral blood leukocytes.
CC Weakly expressed in colon, thymus and small intestine.
CC -1- SIMILARITY: Belongs to the ADIPOR family.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 369.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF151803; AAD34040.1; ALT_FRAME.
DR EMBL; BC001594; AAH01594.1; -.
DR EMBL; BC010743; AAH010743.1; -.
DR MIM; 607945; -.
DR InterPro; IPR004254; HlyIII-related.
DR Pfam; PF03006; UPF0073; 1.
KW Fatty acid metabolism; Lipid metabolism; Receptor; Transmembrane.
FT DOMAIN 1 136
FT TRANSMEM 137 157
FT DOMAIN 158 170
FT TRANSMEM 171 191
FT DOMAIN 192 209
FT TRANSMEM 210 230
FT DOMAIN 231 235
FT TRANSMEM 236 256
FT DOMAIN 257 267
FT TRANSMEM 268 288
FT DOMAIN 289 296
FT TRANSMEM 297 317
FT DOMAIN 318 331
FT TRANSMEM 332 352
FT DOMAIN 353 375
FT SEQUENCE 375 AA; 42615 MW; 1CC0300A7D178E80 CRC64;
SQ
Query Match 76.5%; Score 26; DB 1; Length 375;
Best Local Similarity 50.0%; Pred. No. 44;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy 2 GVXTXIRP 9
Db 152 GILTMLRP 159
ADRI_MOUSE STANDARD; PRT; 375 AA.
AC Q91VHL; Q9CZAO;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Adiponectin receptor protein 1.
GN ADIPOR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

```

Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtauki S.,  
 Hayaizaki Y.,  
 "Functional annotation of a full-length mouse cDNA collection."  
 Nature 409:685-690(2001).  
 (2)  
 SEQUENCE FROM N.A.  
 TISSUE=Breast tumor;  
 MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullah S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."  
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 (3)  
 FUNCTION, AND TISSUE SPECIFICITY.  
 RP MEDLINE=22687101; PubMed=12802337;  
 RA Sugiyama T., Miyagishi M., Hara K., Tsunoda M., Murakami K.,  
 RA Ohteki T., Uchida S., Takekawa S., Waki H., Tsuno N.H., Shibata Y.,  
 RA Terauchi Y., Froguel P., Tobe K., Koyasu S., Taira K., Kitamura T.,  
 RA Shimizu T., Nagai R., Kadowaki T.;  
 "Cloning of adiponectin receptors that mediate antidiabetic metabolic  
 effects."  
 Nature 423:762-769(2003).  
 CC -!- FUNCTION: Receptor for globular and full-length adiponectin  
 (APM1), an essential hormone secreted by adipocytes that acts as  
 an antidiabetic. Probably involved in metabolic pathways that  
 regulate lipid metabolism such as fatty acid oxidation. Mediates  
 increased AMPK, PPARA ligand activity, fatty acid oxidation and  
 glucose uptake by adiponectin. Has some high-affinity receptor for  
 globular adiponectin but low-affinity receptor for full-length  
 adiponectin.  
 CC -!- SUBUNIT: May form homo and heteromultimers (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized to the  
 cell membrane and intracellular organelles (By similarity).  
 CC -!- TISSUE SPECIFICITY: Widely expressed. Expressed in heart, kidney,  
 liver, lung, skeletal muscle and spleen. Weakly expressed in brain  
 and testis.  
 CC -!- SIMILARITY: Belongs to the ADIPOR family.  
 CC -!- CAUTION: Ref1 sequence differs from that shown due to a stop  
 codon in position 353 which was translated as Tyr to extend the  
 sequence and to a frameshift in position 375, which abolishes the  
 stop codon.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AK012847; BAB28509.1; ALT\_SEQ.  
 DR EMBL; BC014875; AAH14875.1;  
 DR MGD; MGI:191924; 281003111Rik.  
 DR InterPro; IPR004254; HlyVII-related.  
 DR Pfam; PF03006; UPR0073; 1.  
 KW Fatty acid metabolism; Lipid metabolism; Receptor; Transmembrane.  
 FT DOMAIN 1 136 CYTOPLASMIC (POTENTIAL).

TRANSMEM 137 157  
 FT DOMAIN 158 170 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 171 191  
 FT DOMAIN 192 209  
 FT TRANSMEM 210 230  
 FT DOMAIN 231 235  
 FT TRANSMEM 236 256  
 FT TRANSMEM 257 267  
 FT TRANSMEM 268 288  
 FT TRANSMEM 289 296  
 FT TRANSMEM 297 317  
 FT TRANSMEM 318 331  
 FT TRANSMEM 332 352  
 FT TRANSMEM 353 375  
 SQ SEQUENCE 375 AA; 42366 MW; 0E72F81B5E9938CE CRC64;  
 Query Match 76.5%; Score 26; DB 1; Length 375;  
 Best Local Similarity 50.0%; Pred. No. 44;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVYXIRP 9  
 : : : :  
 Db 152 GILTMLRP 159  
 : : : :  
 RESULT 13  
 YMN3\_PODAN  
 ID YMN3\_PODAN STANDARD; PRT; 433 AA.  
 AC P15563;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 49.1 kDa protein in ND3 intron.  
 OS Podospora anserina.  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.  
 OC NCBI\_TaxID=5145;  
 OX [1]  
 RN [1] SEQUENCE FROM N.A.  
 RC STRAINS, and A;  
 RX MEDLINE=89125610; PubMed=2975708;  
 RA Cummings D.J., Domenico J.M.;  
 "sequence analysis of mitochondrial DNA from Podospora anserina."  
 RT Pervasiveness of a class I intron in three separate genes."  
 RL J. Mol. Biol. 204:815-839(1988).  
 RN [2]  
 RC COMPLETE GENOME.  
 RP STRAIN=8; PubMed=2357736;  
 RX MEDLINE=90291512; PubMed=2357736;  
 RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;  
 "The complete DNA sequence of the mitochondrial genome of Podospora  
 anserina."  
 RL Curr. Genet. 17:375-402(1990).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X55026; CAA38767.2;  
 DR EMBL; X14485; -; NOT\_ANNOTATED\_CDS.  
 DR FIR; S05654; S05654.  
 DR InterPro; IPR001982; Endonuc\_LAG/HNH.  
 DR Pfam; PF00961; LAGLIDAG 1; 2.  
 KW Hypothetical protein; Mitochondrion.  
 SQ SEQUENCE 433 AA; 49089 MW; 75139328A8B477D0 CRC64;  
 Query Match 76.5%; Score 26; DB 1; Length 433;  
 Best Local Similarity 50.0%; Pred. No. 50;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
| : |||

Db 284 GINTILRP 291

RESULT 14  
VL2 HPV19  
ID - VL2 HPV19 STANDARD; PRT; 520 AA.  
AC P36752;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Minor capsid protein L2.  
GN L2.  
OS Human papillomavirus type 19.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10608;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; PubMed=8205838;  
RA Delius H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X74470; CAA52522.1; -;  
DR PIR; S36489; S36489.  
DR InterPro; IPR000784; Late L2.  
DR Pfam; PF00513; Late protein\_L2; 1.  
KW Coat protein; Late protein.  
SQ SEQUENCE 520 AA; 57019 MW; F428B9097A881AE1 CRC64;  
  
Query Match 76.5%; Score 26; DB 1; Length 520;  
Best Local Similarity 62.5%; Pred. NO. 60;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
| : |||

Db 81 GTATVIRP 88

RESULT 15  
RYR2 HUMAN  
ID - RYR2 HUMAN STANDARD; PRT; 4967 AA.  
AC Q92736; Q15411;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RYR2)  
DE (RYR-2) (Cardiac muscle ryanodine receptor-calcium release channel)  
DE (hRYR-2).  
GN RYR2.  
OS Homo sapiens (Human).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CC NCBI\_TaxID=9606;  
CC [1]  
RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RP TISSUE=Heart muscle;  
RC MEDLINE=96404895; PubMed=8809036;  
RX Tunwell R.E.A.; Wickenden C.; Bertrand B.M.A.; Shevchenko V.I.;  
RA Walsh M.B.; Allen P.D.; Lai F.A.;  
RT "The human cardiac muscle ryanodine receptor-calcium release channel:

RT identification, primary structure and topological analysis.";  
RL Biochem. J. 318:477-487(1996).  
RN [2]  
RP SEQUENCE OF 9-87 AND 533-681 FROM N.A., DEVELOPMENTAL STAGE, AND  
RP INDUCTION.  
RC TISSUE=Heart muscle, and Myometrium;  
RX MEDLINE=97220386; PubMed=9148749;  
RA Avad S.S., Lamb H.K., Morgan J.M., Dunlop W., Gillespie J.I.;  
RT "Differential expression of ryanodine receptor RyR2 mRNA in the non-  
RT pregnant and pregnant human myometrium.";  
RL Biochem. J. 322:777-783(1997).  
RN [3]  
RP SEQUENCE OF 4292-4479 FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=Cerebellum, and Hippocampus;  
RX MEDLINE=98268728; PubMed=9607712;  
RA Martin C., Chapman K.E., Seckl J.R., Ahlley R.H.;  
RT "Partial cloning and differential expression of ryanodine  
RT receptor/calcium-release channel genes in human tissues including the  
RL hippocampus and cerebellum.";  
RL Neuroscience 85:205-216(1998).  
RN [4]  
RP VARIANTS VTSIP LEU-2246; SER-2474; LYS-4104 AND CYS-4497.  
RX PubMed=11208676;  
RA Priori S.G., Napolitano C., Tiso N., Memmi M., Vignati G., Bloise R.,  
RA Sorrentino V.V., Danieli G.A.;  
RT "Mutations in the cardiac ryanodine receptor gene (hRYR2) underlie  
RT catecholaminergic polymorphic ventricular tachycardia.";  
RL Circulation 103:196-200(2001).  
RN [5]  
RP VARIANTS FPVT SER-2328; ARG-4201 AND PHE-4653, AND VARIANT ARG-2958.  
RX MEDLINE=1112862; PubMed=11157710;  
RA Laitinen P.J., Brown K.M., Piippo K., Swan H., Devaney J.M.,  
RA Brahmabhatt B., Donarum E.A., Marino M., Tiso N., Viitasalo M.,  
RA Toivonen L., Stephan D.A., Kortula K.;  
RT "Mutations of the cardiac ryanodine receptor (RYR2) gene in familial  
RT polymorphic ventricular tachycardia.";  
RL Circulation 103:485-490(2001).  
RN [6]  
RP VARIANTS ARVD2 PRO-433 AND ILE-2386.  
RX MEDLINE=21096894; PubMed=11159936;  
RA Tiso N., Stephan D.A., Nava A., Bagattin A., Devaney J.M., Stanchi F.,  
RA Lardet G., Brahmabhatt B., Brown K., Baucé B., Muriago M., Basso C.,  
RA Thiéne G., Danieli G.A., Rampazzo A.;  
RT "Identification of mutations in the cardiac ryanodine receptor gene in  
RT families affected with arrhythmic right ventricular  
RT cardiomyopathy type 2 (ARVD2).";  
RL Hum. Mol. Genet. 10:189-194(2001).  
CC -!- FUNCTION: Communication between transverse-tubules and  
CC sarcoplasmic reticulum. Contraction of cardiac muscle is triggered  
CC by release of calcium ions from SR following depolarization of T-  
CC tubules (By similarity).  
CC -!- SUBUNIT: Homotrimer (Potential).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q92736-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q92736-2; Sequence=VSP\_005953;  
CC TISSUE SPECIFICITY: Heart muscle, brain (cerebellum and  
CC hippocampus) and placenta.  
CC -!- DEVELOPMENTAL STAGE: Expressed in myometrium during pregnancy.  
CC -!- INDUCTION: By TGF-beta.  
CC -!- DISEASE: Defects in RYR2 are the cause of arrhythmic right  
CC ventricular cardiomyopathy type 2 (ARVD2) [MIM:600996]. ARVD2 is  
CC an autosomal dominant disease characterized by partial  
CC degeneration of the myocardium of the right ventricle, electrical  
CC instability, and sudden death. It is clinically defined by  
CC electrocardiographic and angiographic criteria; pathologic  
CC findings, replacement of ventricular myocardium with fatty and  
CC fibrous elements, preferentially involve the right ventricular  
CC free wall.  
CC -!- DISEASE: Defects in RYR2 are the cause of an autosomal dominant

form of stress-induced polymorphic ventricular tachycardia (VTsIP) [MIM:604772]; also known as catecholaminergic polymorphic ventricular tachycardia. VTsIP is a genetic arrhythmic disorder characterized by stress-induced, bidirectional ventricular tachycardia that may degenerate into cardiac arrest and cause sudden death.

-/- DISEASE: Defects in RYR2 are a cause of familial polymorphic ventricular tachycardia (FPVT) [MIM:192605]. FPVT is an autosomal-dominant, inherited disease with a relatively early onset and a mortality rate of approximately 30% by the age of 30 years. Phenotypically, it is characterized by salvos of bidirectional and polymorphic ventricular tachycardias in response to vigorous exercise, with no structural evidence of myocardial disease.

-/- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-/- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-/- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-/- SIMILARITY: Belongs to the ryanodine receptor family.

-/- SIMILARITY: Contains 5 MIR domains.

-/- SIMILARITY: Contains 3 SPRY domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X93330; CAA66975.1; --  
 EMBL; Y08216; CAA63935.1; --  
 EMBL; J91869; CAA62975.1; --  
 EMBL; AJ002511; CAA05502.1; --  
 PIR; S72269; S72269.  
 Genew; HGNC:10484; RYR2.  
 MIM; 180902; --  
 MIM; 600996; --  
 MIM; 604772; --  
 MIM; 192605; --  
 GO; GO:0005886; C:plasma membrane; TAS.  
 GO; GO:0015278; F:calcium-release channel activity; TAS.  
 GO; GO:0006936; P:muscle contraction; TAS.  
 GO; GO:0008016; P:regulation of heart rate; TAS.  
 InterPro; IPR000699; Ca-rel\_channel.  
 InterPro; IPR001682; Ca/Na\_Pore.  
 InterPro; IPR002048; EF-hand.  
 InterPro; IPR005821; Ion\_trans.  
 InterPro; IPR003608; MIR.  
 InterPro; IPR001215; Ryanodn\_receptor.  
 InterPro; IPR003032; RYR.  
 InterPro; IPR003877; SPRY\_receptor.  
 Pfam; PF00036; ehand; 2.  
 Pfam; PF00520; Ion\_trans; 1.  
 Pfam; PF02815; MIR; 4.  
 Pfam; PF01365; RYDR\_ITPR; 2.  
 Pfam; PF02026; RYR; 4.  
 Pfam; PF00632; SPRY; 3.  
 PRINTS; PR00795; RYANODINER.  
 SMART; SM00054; EFH; 2.  
 SMART; SM00472; MIR; 4.  
 SMART; SM00449; SPRY; 3.  
 PROSITE; PS00919; MIR; 5.  
 Ionic channel; Receptor; Calcium channel; Calmodulin-binding; Repeat; Transmembrane; Glycoprotein; Phosphorylation; Alternative splicing; Polymorphism; Disease mutation; Cardiomyopathy; Multigene family.  
 DOMAIN 1 3089 M' (POTENTIAL).  
 TRANSMEM 3090 3109 M' (POTENTIAL).

FT	TRANSMEM	3153	3171	M'' (POTENTIAL).
FT	TRANSMEM	3940	3959	M1 (POTENTIAL).
FT	TRANSMEM	3978	3995	M2 (POTENTIAL).
FT	TRANSMEM	4233	4256	M3 (POTENTIAL).
FT	TRANSMEM	4294	4314	M4 (POTENTIAL).
FT	TRANSMEM	4500	4520	M5 (POTENTIAL).
FT	TRANSMEM	4578	4600	M6 (POTENTIAL).
FT	TRANSMEM	4720	4740	M7 (POTENTIAL).
FT	TRANSMEM	4768	4786	M8 (POTENTIAL).
FT	TRANSMEM	4810	4827	M9 (POTENTIAL).
FT	TRANSMEM	4845	4867	M10 (POTENTIAL).
FT	DOMAIN	110	165	MIR 1.
FT	DOMAIN	172	217	MIR 2.
FT	DOMAIN	225	280	MIR 3.
FT	DOMAIN	286	343	MIR 4.
FT	DOMAIN	351	408	MIR 5.
FT	DOMAIN	670	808	SPRY 1.
FT	DOMAIN	1098	1221	SPRY 2.
FT	DOMAIN	1423	1561	SPRY 3.
FT	DOMAIN	4414	4455	GLU-RICH (ACIDIC).
FT	DOMAIN	853	2925	4 X APPROXIMATE REPEATS.
FT	REPEAT	853	966	1.
FT	REPEAT	967	1080	2.
FT	REPEAT	2892	2810	3.
FT	REPEAT	2812	2925	4.
FT	BINDING	2618	3015	MODULATOR (POTENTIAL).
FT	BINDING	2774	2806	CALMODULIN (POTENTIAL).
FT	BINDING	2876	2897	CALMODULIN (POTENTIAL).
FT	BINDING	2997	3015	CALMODULIN (POTENTIAL).
FT	MOD_RES	2808	2808	PHOSPHORYLATION (BY CAMK) (BY SIMILARITY).
FT	CARBOHYD	3427	3427	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3536	3536	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	3770	3770	N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 76.5%; Score 26; DB 1; Length 4967;

Best Local Similarity 50.0%; Pred. No. 5.3e+02;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

Db 1752 GLSTSLRP 1759

Search completed: April 1, 2004, 17:39:26

Job time : 6.52632 secs

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 28.4211 Seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-4

Perfect score: 34

Sequence: 1 XGVXIXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTRMBL25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp Vertebrate:\*
  - 14: sp Unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	366	12 Q91TT6	Q91tt6 tupaia herp
2	31	91.2	377	16 Q9PPS8	Q9pps8 ureaplasma
3	31	91.2	516	10 Q94D96	Q94d96 oryza sativ
4	30	88.2	96	10 Q880Y8	Q880y8 oryza sativ
5	30	88.2	139	2 Q84CK2	Q84ck2 streptomyc
6	30	88.2	156	10 Q7X6K7	Q7x6k7 oryza sativ
7	30	88.2	2796	2 Q48926	Q48926 mycobacteri
8	29	85.3	185	16 Q9F2Y6	Q9f2y6 streptomyc
9	29	85.3	312	16 Q82HK0	Q82hk0 streptomyc
10	29	85.3	314	16 Q8NLF3	Q8nlf3 corynebacte
11	29	85.3	407	16 Q8F364	Q8f364 leptospira
12	29	85.3	464	16 Q89QR0	Q89qr0 bradyrhizob
13	29	85.3	502	16 Q7WPH7	Q7wph7 bordetella
14	29	85.3	619	2 Q8RLI7	Q8rll7 salmonella
15	29	85.3	5035	6 Q29104	Q29104 sus scrofa
16	29	85.3	5035	6 Q29105	Q29105 sus scrofa

Q80x16 mus musculus  
Q9hs92 halobacteri  
O07538 bacillus su  
P96566 amycolatops  
Q91108 oryza sativ  
Q8r946 thermoaer  
Q22659 caenorhabdi  
Q85716 mycobacteri  
Q88df6 mycoplasma  
Q82br2 streptomyc  
Q9kdq9 bacillus ha  
Q9cr40 mus musculu  
Q87h35 vibrio para  
Q8iue3 homo sapien  
Q91214 streptomyc  
Q8zx98 bacterioph  
Q93919 human cytom  
Q86312 unidentifie  
Q8zwe4 pyrobaculum  
Q8u0f4 pyrococcus  
Q8pgv5 xanthomonas  
Q88y76 lactobacill  
Q8f620 rhizobium m  
Q92rz1 rhizobium m  
Q97da0 clostridium  
Q7zy04 xenopus lae  
Q9asp7 caulobacter  
Q8luj4 bacillus an  
Q8vpp4 micrococcus

ALIGNMENTS

RESULT 1

Q91TT6 PRELIMINARY; PRT; 366 AA.  
ID Q91TT6  
AC Q91TT6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE T22.3.  
OS Tupaia herpesvirus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae.  
OX NCBI\_TaxID=10397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RX MEDLINE=21211637; PubMed=11312357;  
RA Bahr U., Darai G.;  
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree  
Shrew) Herpesvirus.";  
RL J. Virol. 75:4854-4870(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RA Darai G., Bahr U.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF281817; AAK57051.1; -.  
SQ SEQUENCE 366 AA; 42795 MW; 757ED66A38D827FE CRC64;

Query Match 91.2%; Score 31; DB 12; Length 366;

Best Local Similarity 75.0%; Pred. No. 21;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

Db 162 GVSTSIIRP 169

RESULT 2

Q9PPS8

ID Q9PPS8 PRELIMINARY; PRT; 377 AA.  
AC Q9PPS8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Oligopeptide transport system permease protein.  
GN O9PB OR U5562.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Serovar 3;  
RC MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762(2000).  
DR EMBL; AE002154; AAF30975.1; --  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD.transp.  
DR Pfam; PF00528; BPD.transp; 1.  
KW Complete proteome.  
SQ SEQUENCE 377 AA; 41923 MW; 391626240B625A6F CRC64;

Query Match 91.2%; Score 31; DB 16; Length 377;  
Best Local Similarity 75.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
||| |||  
Db 146 GVSTSRP 153

#### RESULT 3

ID Q94D96 PRELIMINARY; PRT; 516 AA.  
AC Q94D96;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE P0439E11.9 protein.  
GN P0439E11.9.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=cv. Nipponbare;  
RC Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:P0439E11.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003315; BA61259.1; --  
DR Gramene; Q94D96; --  
DR GO; GO:0005229; F:sugar binding; IEA.  
DR InterPro; IPR001480; B:lectin.  
DR Pfam; PF01453; Agglutinin; 1.  
DR SMART; SM00108; B:lectin; 1.  
SQ SEQUENCE 516 AA; 55483 MW; CB8B8B9B56377700 CRC64;

Query Match 91.2%; Score 31; DB 10; Length 516;  
Best Local Similarity 75.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
||| |||  
Db 349 GVTTAIRP 356

#### RESULT 4

ID Q850Y8 PRELIMINARY; PRT; 96 AA.  
AC Q850Y8;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE P0485B12.29 protein (P0031D02.2 protein).  
GN P0485B12.29 OR P0031D02.2.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=39947;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=cv. Nipponbare;  
RC Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC  
clone:P0485B12.2";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003348; BA86498.1; --  
DR EMBL; AP003230; BA893174.1; --  
DR Gramene; Q850Y8; --  
SQ SEQUENCE 96 AA; 10908 MW; 2F93BF8513818BF8 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 96;  
Best Local Similarity 62.5%; Pred. No. 9.7;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
||| |||  
Db 67 GVTTSVRP 74

#### RESULT 5

ID Q84CK2 PRELIMINARY; PRT; 139 AA.  
AC Q84CK2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE KINUL.  
GN KINUL.  
OS Streptomyces murayamaensis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
OX NCBI\_TaxID=224537;  
[1]  
RN SEQUENCE FROM N.A.  
RP Yim X.H., Mahadevan B., Grochowski L., Proteau P.J.;  
RT "Molecular cloning and sequence of the kinamycin angucycline type II  
polyketide synthase gene cluster from Streptomyces murayamaensis.";  
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY228175; AAO65341.1; --  
DR InterPro; IPR009002; FMN binding.  
SQ SEQUENCE 139 AA; 14819 MW; 65A27123A80349E4 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 139;  
Best Local Similarity 75.0%; Pred. No. 14;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
||| |||



Db 28 GVLTTIRP 35

RESULT 6

Q7X6K7 PRELIMINARY; PRT; 156 AA.

ID Q7X6K7

AC Q7X6K7

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE OSJNBA001819.23 protein (OSJNBA0034E24.1 protein).

GN OSJNBA001819.23 OR OSJNBA0034E24.1.

OS Oryza sativa (rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzoideae; Oryza.

OX NCBI\_TaxID=4530;

RN [1]

RP SEQUENCE FROM N.A.

RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,

RA Liu Y.L., Wu J., Yu Z., Chen L., Fan D.L., Meng Q.J., Zhang L.,

RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,

RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Yang K., Zhou B., Chen Z.H.,

RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,

RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,

RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,

RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,

RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL662956; CAE04456.1; -

DR EMBL; AL662957; CAE05907.1; -

SQ SEQUENCE 156 AA; 17088 MW; C8B47B96DFB7BD02 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 156;

Best Local Similarity 62.5%; Pred. No. 16;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 30 GVTSVRP 37

RESULT 7

Q48926 PRELIMINARY; PRT; 2796 AA.

ID Q48926

AC Q48926

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Fatty acid synthase.

GN PAS.

OS Mycobacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1765;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=BCG-Pasteur;

RA MEDLINE=96200863; PubMed=8621098;

RA Fernandes N.D., Kolattukudy P.S.;

RT "Cloning, sequencing and characterization of a fatty acid synthase-

RT encoding gene from Mycobacterium tuberculosis var. bovis BCG.";

RL Gene 170:95-99(1996).

DR EMBL; U36763; AAB03809.1; -

DR PIR; JCA743; JCA743.

DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.

DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR001527; Ac trans.

DR InterPro; IPR003965; Fatty\_acid\_synth.

DR InterPro; IPR000794; Ketoacyl\_synth.

DR InterPro; IPR002539; Maoc\_dehydratas.

DR Pfam; PF00698; Acyl\_transf; 1.

DR Pfam; PF00109; ketoacyl-synt; 1.

DR Pfam; PF02801; ketoacyl-synt; C; 1.

DR Pfam; PF01575; Maoc\_dehydratas; 1.

DR PRINTS; PRO4483; FASINIHASE.

DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 1.

SQ SEQUENCE 2796 AA; 303675 MW; 47B87169126D2504 CRC64;

Query Match 88.3%; Score 30; DB 2; Length 2796;

Best Local Similarity 75.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 2704 GVTSIRP 2711

RESULT 8

Q9F216 PRELIMINARY; PRT; 185 AA.

ID Q9F216

AC Q9F216

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative integral membrane protein.

GN SC04361 OR SCD19.16.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;

RX MEDLINE=21996410; PubMed=12000953;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.P., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

RT coelicolor A3(2).";

RL Nature 417:141-147(2002).

DR EMBL; AL939119; CAC08317.1; -

KW Complete proteome.

SQ SEQUENCE 185 AA; 19332 MW; E2FE76A5D21296FD CRC64;

Query Match 85.3%; Score 29; DB 16; Length 185;

Best Local Similarity 75.0%; Pred. No. 34;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 73 GVTSIRP 80

RESULT 9

Q82HKO PRELIMINARY; PRT; 312 AA.

ID Q82HKO

AC Q82HKO

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative secreted substrate-binding protein.

GN SAV3508.

OS Streptomyces avermitilis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

```

OX NCBI_TaxID=33903;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608305; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005035; BAC71220.1; -.
DR InterPro; IPR001092; HLH_basic.
DR InterPro; IPR007210; OpuAC.
DR Pfam; PF04069; OpuAC; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 312 AA; 32672 MW; CADA8070DA5C2210 CRC64;

Query Match 85.3%; Score 29; DB 16; Length 312;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 71 GVRTTVP 78

RESULT 10
Q8NLF3 PRELIMINARY; PRT; 314 AA.
AC Q8NLF3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Cgl2889.
GN CGL2889.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005283; BAC00383.1; -.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006016; Usp_dom.
DR Pfam; PF00582; Usp; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 314 AA; 34512 MW; B2423836E2E4DBDC CRC64;

Query Match 85.3%; Score 29; DB 16; Length 314;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 46 GITTIFP 53

```

```

RESULT 11
Q8F364 PRELIMINARY; PRT; 407 AA.
AC Q8F364;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capb protein, Mur ligase family.
GN CAPB OR LA2547.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=1173;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=56501 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011422; AAN49746.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0045227; P:capsule polysaccharide biosynthesis; IEA.
DR InterPro; IPR008337; CapB.
DR InterPro; IPR001005; Wyt DNA binding.
DR PRINTS; PR01758; CAPSULEPROTB.
DR PROSITE; PS00037; MIB_1; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 407 AA; 46759 MW; AF12FAB606746B4B CRC64;

Query Match 85.3%; Score 29; DB 16; Length 407;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 132 GVITNIRP 139

RESULT 12
Q89QRO PRELIMINARY; PRT; 464 AA.
AC Q89QRO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Succinyl-diaminopimelate desuccinylase.
GN DAPE OR BLR3064.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005946; BAC48329.1; -.
DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase M20.
DR Pfam; PF01546; Peptidase_M20; 1.
KW Complete proteome.
SQ SEQUENCE 464 AA; 50583 MW; B12C32D8F9432CA1 CRC64;

```

```
Query Match      85.3%; Score 29; DB 16; Length 464;
Best Local Similarity 62.5%; Pred. No. 85;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      58 GVATEVRP 65

RESULT 13
Q7WPH7 PRELIMINARY; PRT; 502 AA.
AC Q7WPH7;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Probable GntR-family transcriptional regulator.
GN B30707
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cordero-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsis K.,
RA Leather S., Mout R., Norbertzak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX40439; CA31207.1; -.
DR Complete proteome.
SQ SEQUENCE 502 AA; 54866 MW; 19C15B8F2F34EB3D CRC64;

Query Match      85.3%; Score 29; DB 16; Length 502;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      424 GIATAVRP 431

RESULT 14
Q8RL7 PRELIMINARY; PRT; 619 AA.
AC Q8RL7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
OS Salmonella enteritidis.
CG Plasmid pP.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=592;
RN [1]
RP SEQUENCE FROM N.A.
RA Rychlik I., Gregorova D., Doskocilova J., Sebkova A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY079199; AL87010.1; -.
DR GO; GO:0946821; C:extrachromosomal DNA; IEA.
DR GO; GO:0009291; P:unidirectional conjugation; IEA.
DR InterPro; IPR005053; MobA_MobL.
```

```
DR Pfam; PF03389; MobA_MobL; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 619 AA; 71857 MW; BF15FF501426DAF5 CRC64;

Query Match      85.3%; Score 29; DB 2; Length 619;
Best Local Similarity 75.0%; Pred. No. 11e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      528 GVKTEIRP 535

RESULT 15
Q29104 PRELIMINARY; PRT; 5035 AA.
AC Q29104;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ryanodine receptor.
GN RyR1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Belgisch Landras/Pietrain; TISSUE=Skeletal muscle;
RX MEDLINE=91320118; PubMed=1862346;
RA Fujii J., Otsu K., Zorzato F., de Leon S., Khanna V.K., Weiler J.E.,
RA O'Brien P.J., MacLennan D.H.;
RT "Identification of a mutation in porcine ryanodine receptor associated
RT with malignant hyperthermia.";
RL Science 253:448-451(1991).
DR EMBL; M91451; AAA3118.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005262; F:calcium channel activity; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006816; P:calcium ion transport; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR000699; Ca-rel_channel.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion_trans.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR01215; Ryanodn_receptor.
DR InterPro; IPR003032; RyR.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITPR; 2.
DR Pfam; PF02026; RYR; 4.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Ionic channel; Receptor; Transmembrane.
SQ SEQUENCE 5035 AA; 565089 MW; 66B0E1AA590636EF CRC64;

Query Match      85.3%; Score 29; DB 6; Length 5035;
Best Local Similarity 62.5%; Pred. No. 9.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      1767 GVTISLRP 1774

Search completed: April 1, 2004, 17:38:41
Job time : 29.4211 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 43.1053 Seconds  
(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-5

Perfect score: 34

Sequence: 1 XGVXTXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	88.2	8	7	ADD31284 Angiogene
2	30	88.2	8	7	AD83635 Antiangi
3	30	88.2	8	7	AD83635 Antiangi
4	30	88.2	9	3	AA67983 Antiangi
5	30	88.2	9	4	AAU15042 Antiangi
6	30	88.2	9	4	AAU15043 Antiangi
7	30	88.2	9	4	AAU15027 Antiangi
8	30	88.2	9	6	AAO26633 Anti-angi
9	30	88.2	10	4	AAU14987 Antiangi
10	30	88.2	10	4	AAU14993 Antiangi
11	30	88.2	139	6	ABP57616 S. muraya
12	29	85.3	2796	6	ABU79133 Mycobacte
13	29	85.3	8	7	ADD31274 Angiogene
14	29	85.3	8	7	ADD31265 Angiogene
15	29	85.3	8	7	ADD31264 Angiogene
16	29	85.3	8	7	ADD31266 Angiogene
17	29	85.3	8	7	ADD31294 Angiogene
18	29	85.3	8	7	ADD31288 Angiogene
19	29	85.3	8	7	ADD31273 Angiogene
20	29	85.3	8	7	ADD31269 Angiogene
21	29	85.3	8	7	ADD31267 Angiogene
22	29	85.3	8	7	ADD31282 Angiogene
23	29	85.3	8	7	AD83640 Antiangi
24	29	85.3	8	7	AD83694 Antiangi
25	29	85.3	8	7	AD83632 Antiangi
26	29	85.3	8	7	AD83636 Antiangi

26	29	85.3	8	7	AD83629 Antiangi
27	29	85.3	8	7	AD83628 Antiangi
28	29	85.3	8	7	AD83641 Antiangi
29	29	85.3	8	7	AD83631 Antiangi
30	29	85.3	8	7	AD83691 Antiangi
31	29	85.3	9	3	AAV67973 Antiangi
32	29	85.3	9	3	AAV67978 Antiangi
33	29	85.3	9	3	AAV67992 Antiangi
34	29	85.3	9	3	AAV67975 Antiangi
35	29	85.3	9	3	AAV67986 Antiangi
36	29	85.3	9	3	AAV67996 Antiangi
37	29	85.3	9	3	AAV67997 Antiangi
38	29	85.3	9	3	AAV67979 Antiangi
39	29	85.3	9	3	AAV67995 Antiangi
40	29	85.3	9	3	AAV67985 Antiangi
41	29	85.3	9	3	AAV67976 Antiangi
42	29	85.3	9	3	AAV67989 Antiangi
43	29	85.3	9	3	AAV67974 Antiangi
44	29	85.3	9	3	AAV67981 Antiangi
45	29	85.3	9	3	AAV67982 Antiangi

## ALIGNMENTS

RESULT 1  
ADD31284  
ID ADD31284 standard; peptide; 8 AA.  
XX  
AC ADD31284;  
XX

DT 15-JAN-2004 (first entry)  
XX  
DE Angiogenesis inhibiting peptide #59.  
XX

KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
OS Synthetic.  
XX

Key Location/Qualifiers  
FT Modified-site 1 /label= OTHER  
FT Modified-site 3 /note= "OTHER= N-acetyl"  
FT Modified-site 8 /label= OTHER  
FT Modified-site 8 /note= "OTHER= D-form residue, aile (not defined)"  
FT Modified-site 8 /note= "NHCH2CH3"

US2003109455-A1.  
12-JUN-2003.  
30-OCT-2002; 2002US-00283550.  
31-OCT-2001; 2001US-0335017P.  
(HAVI/) HAVIV F.  
(BRAD/) BRADLEY M F.  
Haviv P, Bradley MF;  
WPI; 2003-843101/78.  
New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
angiogenesis, treating cancer in mammal, or for prevention of other  
diseases such as autoimmune diseases.  
Claim 13; Page 25; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.

XX Sequence 8 AA;

Query Match 88.2%; Score 30; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVXTXIRP 9  
 |||||  
 Db 1 GVXTSIRP 8

RESULT 2  
 ADE83635  
 ID ADE83635 standard; peptide; 8 AA.

XX ADE83635;

XX 29-JAN-2004 (first entry)

XX Antiangiogenic peptide.

XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiac; vulnerary;  
 KW antinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.

XX Synthetic.

XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "acetylated"

FT Modified-site 3 /note= "D-form residue alloisoleucyl"

FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"

XX WO2003037268-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-US034811.

XX 31-OCT-2001; 2001US-00000681.

XX 04-OCT-2002; 2002US-00263812.

XX (ABBO ) ABBOTT LAB.

XX Haviv F, Bradley MF;

XX WPI; 2003-617886/58.

XX New hepta-, octa- and nona-peptide compounds used for treating e.g.  
 PT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic  
 PT retinopathy.

XX Claim 13; Page 46; 51pp; English.

CC The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antiatherosclerotic,  
 CC cardiac, vulnerary, antiinflammatory, antiulcer and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumours. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathological  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 88.2%; Score 30; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 GVXTXIRP 9  
 |||||  
 Db 1 GVXTSIRP 8

RESULT 3

AAV67983

ID AAV67983 standard; peptide; 9 AA.

XX AAV67983;

XX 11-APR-2000 (first entry)

XX Antiangiogenic peptide #12.

XX Antiangiogenic; antiarthritic; cytostatic; antipsoriatic; arthritis;  
 KW angiogenesis; inhibitor; cancer; psoriasis; infection; ocular disease;  
 KW macular degeneration; diabetic retinopathy; tumour metastasis;  
 KW autoimmune disease; neovascularisation; Crohn's disease; birth control;  
 KW cat scratch disease.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= MeGly  
 FT /note= "Sarcosine (methylglycine); acetylated"

FT Misc-difference 4

FT /note= "D form residue"

FT Modified-site 9 /note= "Pro is modified to ProNHCH2CH3"

XX WO9961476-A1.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US011448.

XX 22-MAY-1998; 98US-00083745.

XX 16-FEB-1999; 99US-00250574.

XX 26-MAR-1999; 99US-00277466.

XX (ABBO ) ABBOTT LAB.

XX Henkin J, Haviv F, Bradley MF, Kalvin DM, Schneider AJ;

XX WPI; 2000-072606/06.

XX

PT New anti-angiogenic peptides, used for treating e.g. cancer, arthritis,  
PT psoriasis, or angiogenesis of the eye associated with infection or  
PT surgical intervention, macular degeneration and diabetic retinopathy.  
XX  
XX  
XX Claim 12; Page 75; 223pp; English.  
XX  
XX AAY67972 to AAY67999 represent novel anti-angiogenic peptides. Peptide  
CC from the present invention have the formula: A0-A1-A2-A3-A4-A5-A6-A7-A8-  
CC A9-A10 where A0 is selected from hydrogen or an acyl group; A10 is a  
CC hydroxyl group or an amino acid amide; and A1 to A9 are amino acyl  
CC residues. The peptides are used for anti-angiogenesis therapy in  
CC patients. They are used to treat e.g. cancer, arthritis, psoriasis,  
CC angiogenesis of the eye associated with infection or surgical  
CC intervention, macular degeneration and diabetic retinopathy. They can  
CC prevent tumour metastases. Further uses include treatment and prophylaxis  
CC of autoimmune diseases, various ocular diseases e.g. diabetic  
CC retinopathy, and other abnormal neovascularisation conditions of the eye,  
CC skin diseases e.g. psoriasis, diseases characterised by excessive or  
CC abnormal stimulation of endothelial cells, e.g. Crohn's disease. They can  
CC also be used as a birth control agent, inhibiting ovulation and placental  
CC establishment, and to treat diseases that have angiogenesis as a  
CC pathological consequence e.g. cat scratch disease and are useful to  
CC reduce bleeding by administration prior to surgery, especially for the  
CC treatment of tumours. They can also be used to isolate a receptor from an  
CC endothelial cell  
XX  
XX Sequence 9 AA;

Query Match 88.2%; Score 30; DB 3; Length 9;  
Best Local Similarity 75.8%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
Matches 6; Conservative 0; Mismatches 2;  
QY 2 GVXTXIRP 9  
Db 2 GVTAIRP 9

RESULT 4  
AAU15042  
ID AAU15042 standard; peptide; 9 AA.  
XX  
XX AAU15042;  
XX  
XX 04-DEC-2001 (first entry)  
XX  
DE Antiangiogenic peptide #69 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
XX Modified-site 1  
FT /label= MeGly  
FT /note= "N-methyl-glycine (sarcosine), additionally  
FT modified by N-terminal acetyl"  
FT Modified-site 4  
FT /note= "D-form residue, alioisoleucine"  
FT Modified-site 6  
FT /note= "N-Methyl serine"  
FT Modified-site 9  
FT /note= "Modified by NH-ethyl"

XX WO200138397-A1.  
XX  
XX 31-MAY-2001.  
XX  
XX 22-NOV-2000; 2000WO-US032105.  
XX  
XX 22-NOV-1999; 99US-00447099.

PR 31-OCT-2000; 2000US-00702649.  
XX (ABBO ) ABBOTT LAB.  
FA  
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
PI WPI; 2001-521804/57.  
XX  
DR  
XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
FT diabetic retinopathy.  
XX  
XX Claim 39; Page 89; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides  
CC (AAU1974-AAU1952) which can act as angiogenesis inhibitors. The  
CC antiangiogenic peptides are useful for isolating a receptor from an  
CC endothelial cell. The peptides of the invention are also useful for  
CC treating cancer, arthritis, psoriasis and other skin diseases,  
CC angiogenesis of the eye associated with infection or surgical  
CC intervention and other ocular diseases, cat scratch disease, ulcers,  
CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
CC haemangiomas and capillary action within atherosclerotic plaques,  
CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
CC Osler-Webber syndrome, myocardial angiogenesis, plaque  
CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
CC wound granulation, excessive or abnormal stimulation of endothelial  
CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
CC peptides are also useful as birth control agents by inhibiting ovulation  
CC and to reduce bleeding by administration before surgery. The peptides of  
CC the invention exhibit improved metabolic stability, improved  
CC pharmacokinetics, increased water solubility, and improved oral  
CC availability. The present sequence represents antiangiogenic peptide #69  
XX  
XX Sequence 9 AA;

Query Match 88.2%; Score 30; DB 4; Length 9;  
Best Local Similarity 87.5%; Pred. No. 1.4e+06; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 1;  
QY 2 GVXTXIRP 9  
Db 2 GVXTSIRP 9

RESULT 5  
AAU15043  
ID AAU15043 standard; peptide; 9 AA.  
XX  
XX AAU15043;  
XX  
DT 04-DEC-2001 (first entry)  
XX  
XX Antiangiogenic peptide #70 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
XX Modified-site 1  
FT /label= MeGly  
FT /note= "N-methyl-glycine (sarcosine), additionally  
FT modified by N-terminal acetyl"  
FT Misc-difference 4  
FT /note= "D-form residue"  
FT Modified-site 6  
FT /note= "N-methyl serine"  
FT Modified-site 9

FT XX /note= "Modified by NH-ethyl"

FN WO200138397-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032105.

XX 22-NOV-1999; 99US-00447099.

PR 31-OCT-2000; 2000US-00702649.

XX (ABBO ) ABBOTT LAB.

XX Haviv F, Henkin J, Bradley MF, Kalvin DM;

XX WPI; 2001-521804/57.

XX New N-alkylated peptide compounds useful for treating e.g. cancer,

PT autoimmune diseases, arthritis, psoriasis, macular degeneration and

PT diabetic retinopathy.

XX Claim 39; Page 89; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides

CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The

CC antiangiogenic peptides are useful for isolating a receptor from an

CC endothelial cell. The peptides of the invention are also useful for

CC treating cancer, arthritis, psoriasis and other skin diseases,

CC angiogenesis of the eye associated with infection or surgical

CC intervention and other ocular diseases, cat scratch disease, ulcers,

CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.

CC haemangiomas and capillary action within atherosclerotic plaques,

CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,

CC Osler-Weber syndrome, myocardial angiogenesis, plaque

CC neovascularisation, telangiectasia, haemophilic joint, angiofibroma,

CC wound granulation, excessive or abnormal stimulation of endothelial

CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,

CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic

CC peptides are also useful as birth control agents by inhibiting ovulation

CC and to reduce bleeding by administration before surgery. The peptides of

CC the invention exhibit improved metabolic stability, improved

CC pharmacokinetics, increased water solubility, and improved oral

CC availability. The present sequence represents antiangiogenic peptide #70

XX Sequence 9 AA;

XX Query Match 88.2%; Score 30; DB 4; Length 9;

XX Best Local Similarity 75.0%; Pred. No. 1.4e+06;

XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 2 GVITSIRP 9

RESULT 6

AAU15027

ID AAU15027 standard; peptide; 9 AA.

XX AAU15027;

XX 04-DEC-2001 (first entry)

XX Antiangiogenic peptide #54 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;

XX arthritis; skin disease; ocular disease; diabetic retinopathy;

XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;

XX cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= Megly

FT /note= "N-methyl-glycine (sarcosine), additionally

FT modified by N-terminal acetyl"

FT Misc-difference 4 /note= "D-form residue"

FT Modified-site 6 /note= "N-Methyl alanine"

FT Modified-site 9 /note= "Modified by NH-ethyl"

XX WO200138397-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032105.

XX 22-NOV-1999; 99US-00447099.

PR 31-OCT-2000; 2000US-00702649.

XX (ABBO ) ABBOTT LAB.

XX Haviv F, Henkin J, Bradley MF, Kalvin DM;

XX WPI; 2001-521804/57.

XX New N-alkylated peptide compounds useful for treating e.g. cancer,

PT autoimmune diseases, arthritis, psoriasis, macular degeneration and

PT diabetic retinopathy.

XX Claim 39; Page 89; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides

CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The

CC antiangiogenic peptides are useful for isolating a receptor from an

CC endothelial cell. The peptides of the invention are also useful for

CC treating cancer, arthritis, psoriasis and other skin diseases,

CC angiogenesis of the eye associated with infection or surgical

CC intervention and other ocular diseases, cat scratch disease, ulcers,

CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.

CC haemangiomas and capillary action within atherosclerotic plaques,

CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,

CC Osler-Weber syndrome, myocardial angiogenesis, plaque

CC neovascularisation, telangiectasia, haemophilic joint, angiofibroma,

CC wound granulation, excessive or abnormal stimulation of endothelial

CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,

CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic

CC peptides are also useful as birth control agents by inhibiting ovulation

CC and to reduce bleeding by administration before surgery. The peptides of

CC the invention exhibit improved metabolic stability, improved

CC pharmacokinetics, increased water solubility, and improved oral

CC availability. The present sequence represents antiangiogenic peptide #54

XX Sequence 9 AA;

XX Query Match 88.2%; Score 30; DB 4; Length 9;

XX Best Local Similarity 75.0%; Pred. No. 1.4e+06;

XX Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

DB 2 GVITAIRP 9

RESULT 7

AAO26633

ID AAO26633 standard; peptide; 9 AA.

XX AAO26633;

XX 28-MAR-2003 (first entry)

XX Anti-angiogenic peptide #16.

AC	AAU14987;
XX	
DT	04-DEC-2001 (first entry)
XX	
DE	Antiangiogenic peptide #14 useful for inhibiting angiogenesis.
XX	
KW	Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;
XX	arthritis; skin disease; ocular disease; diabetic retinopathy;
KW	blood vessel disease; atherosclerosis; autoimmune disease; birth control;
KW	cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.
XX	
OS	Synthetic.
XX	
Key	Location/Qualifiers
XX	
FF	Modified-site 1
FT	/label= MeGly
FT	/note= "N-methyl-glycine (sarcosine), additionally
FT	modified by N-terminal acetyl"
FT	.4
FT	Modified-site
FT	/note= "D-form residue, alloisoleucine"
FT	6
FT	Modified-site
FT	/note= "N-Methyl serine"
FT	10
FT	Modified-site
FT	/note= "D-form residue, C-terminal amide"
XX	
PN	WO200138397-A1.
XX	
PD	31-MAY-2001.
XX	
PD	22-NOV-2000; 2000WO-US032105.
XX	
PP	22-NOV-1999; 99US-00447099.
XX	
PR	31-OCT-2000; 2000US-00702649.
XX	
PA	(ABBO ) ABBOTT LAB.
XX	
PI	Haviv F, Henkin J, Bradley MF, Kalvin DM;
XX	
PI	WPI; 2001-521804/57.
XX	
DR	New N-alkylated peptide compounds useful for treating e.g. cancer,
PT	autoimmune diseases, arthritis, psoriasis, macular degeneration and
PT	diabetic retinopathy.
XX	
PS	Claim 38; Page 88; 95pp; English.
XX	
CC	The present invention relates to novel synthetic antiangiogenic peptides
XX	(AAU14974-AAU15052) which can act as angiogenesis inhibitors. The
CC	antiangiogenic peptides are useful for isolating a receptor from an
CC	endothelial cell. The peptides of the invention are also useful for
CC	treating cancer, arthritis, psoriasis and other skin diseases,
CC	angiogenesis of the eye associated with infection or surgical
CC	intervention and other ocular diseases, cat scratch disease, ulcers,
CC	macular degeneration and diabetic retinopathy, blood vessel diseases e.g.
CC	haemangiomas and capillary action within atherosclerotic plaques,
CC	autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,
CC	Ogler-Webber syndrome, myocardial angiogenesis, plaque
CC	neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
CC	wound granulation, excessive or abnormal stimulation of endothelial
CC	cells, including intestinal adhesions, Crohn's disease, atherosclerosis,
CC	scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic
CC	peptides are also useful as birth control agents by inhibiting ovulation
CC	and to reduce bleeding by administration before surgery. The peptides of
CC	the invention exhibit improved metabolic stability, improved
CC	pharmacokinetics, increased water solubility, and improved oral
CC	availability. The present sequence represents antiangiogenic peptide #14
XX	
XX	Sequence 10 AA;
XX	
XX	Query Match 88.2%; Score 30; DB 4; Length 10;
XX	Best Local Similarity 87.5%; Pred. No. 0.83;
XX	Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



CC	pharmacokinetics, increased water solubility, and improved oral
CC	availability. The present sequence represents antiangiogenic peptide #20
XX	
SQ	Sequence 10 AA;
Query Match	88.2%; Score 30; DB 4; Length 10;
Best Local Similarity	75.0%; Pred. NO. 0.83;
Matches	6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	2 GVXTXIRP 9 
Db	2 GVITSIRP 9 
RESULT 10	
ABP57616	
ID	ABP57616 standard; protein; 139 AA.
XX	
AC	ABP57616;
XX	
DT	29-APR-2003 (first entry)
XX	
DE	S. murayamaensis ATCC 21414 kinamycin protein SEQ ID NO:17.
XX	
KW	Streptomyces murayamaensis ATCC 21414; Streptomyces murayamaensis;
KW	glycosylated kinamycin; kinamycin; type II polyketide; polyketide;
KW	antibacterial; cytostatic; infection; antibiotic; antitumour;
KW	electrophilic azo-coupling agent.
XX	
OS	Streptomyces murayamaensis.
XX	
PN	WO2003002066-A2.
XX	
PD	09-JAN-2003.
XX	
PP	27-JUN-2002; 2002WO-US020719.
XX	
PR	27-JUN-2001; 2001US-0301401P.
XX	
PA	(DIVE-) DIVERSA CORP.
XX	
PI	Short JM, Paradkar A, Varoglu M, Mathur EU;
XX	
DR	WFI; 2003-210195/20.
DR	N-PSDB; ABZ71139.
XX	
PT	New isolated polyketide used e.g. as antibiotic and antitumor agents
PT	comprises kinamycin molecule comprising at least one saccharide group.
XX	
PS	Claim 75; Page 91-92; 119pp; English.
XX	
CC	The present invention describes a polyketide comprising a kinamycin
CC	molecule comprising at least one saccharide group. ABZ71132 to ABZ71163
CC	encode glycosylated kinamycins ABP57609 to ABP57640 isolated from
CC	Streptomyces murayamaensis ATCC 21414. Kinamycins are a class of type II
CC	Polyketides. The kinamycins have antibacterial and cytostatic activities.
CC	They can be used for treating infections as antibiotics and as antitumour
CC	agents, and as electrophilic azo-coupling agents in vitro or in vivo
XX	
SQ	Sequence 139 AA;
Query Match	88.2%; Score 30; DB 6; Length 139;
Best Local Similarity	75.0%; Pred. No. 15;
Matches	6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	2 GVXTXIRP 9 
Db	28 GVLTTIRP 35 
RESULT 11	
ABU79133	
ID	ABU79133 standard; protein; 2796 AA.

XX AC ABU79133;  
 XX DT 18-JUN-2003 (first entry)  
 XX DE Mycobacteria mycolic acid protein.  
 XX KW Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
 XX KW gene therapy; mammalian cell receptor; cytostatic;  
 XX KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
 XX KW tumouricidal immunocyte; antitumour.  
 XX OS Mycobacteria sp.  
 XX PN US2002177551-A1.  
 XX PD 28-NOV-2002.  
 XX PF 30-MAY-2001; 2001US-00870759.  
 XX PR 31-MAY-2000; 2000US-0208128P.  
 XX PX (TERM/) Terman D S.  
 XX PY Terman DS;  
 XX WPI; 2003-361759/34.  
 XX N-PSDB; ACA64735.  
 XX A mammalian cell receptor, useful in the treatment of cancer by binding  
 XX to tumour associated lipids where the binding induces energy or apoptosis  
 XX in T cells and antigen presenting cells.  
 XX Example 2; Page; 167pp; English.  
 XX The invention relates to a mammalian cell receptor, useful in the  
 XX treatment of cancer, which binds to tumour associated lipids and induces  
 XX energy or apoptosis in the T cells and antigen presenting cells (APCs).  
 XX Also included are a mammalian cell useful in the treatment of cancer  
 XX where the receptor which binds tumour associated lipids and induces  
 XX cellular inactivation or death is deleted or functionally deactivated,  
 XX producing (M1) a tumouricidal immunocyte population in vivo in a mammal  
 XX (by allowing tumour associated lipids to contact immunocytes in which  
 XX receptors for immunosuppressive fatty acids, ceramides, glycolipids,  
 XX sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides, or  
 XX sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
 XX deleted), a construct useful in the treatment of cancer comprising a  
 XX superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell  
 XX useful in the treatment of cancer (where an adaptor protein which  
 XX inhibits T cell activation by tumour associated antigens is deleted or  
 XX functionally deactivated), a composition useful in the treatment of  
 XX cancer (comprising a lipid raft conjugated to a superantigen), producing  
 XX (M2) a tumouricidal immunocyte population ex vivo in a mammal (by  
 XX allowing tumour associated lipids to contact immunocytes, in which  
 XX receptors for the lipids are inactivated or deleted to produce a  
 XX tumouricidal immunocyte population, and administering the tumouricidal APC  
 XX activated immunocytes to the host), producing (M3) a tumouricidal APC  
 XX population ex vivo in a mammal (by allowing a tumour associated lipid to  
 XX contact APCs, in which receptors for the tumour associated lipids are  
 XX inactivated or deleted to produce a tumouricidally activated population,  
 XX and administering APCs to the host), producing a tumouricidal T cell  
 XX population ex vivo in a mammal (by allowing a tumour associated lipids to  
 XX contact T cells, in which adaptor proteins, which inhibit T cell  
 XX activation by tumour associated antigens, are deleted or functionally  
 XX deactivated to produce a tumouricidal population of T cells, and  
 XX administering the tumouricidally activated T cells to the host, or  
 XX allowing a superantigen-lipid raft to contact T cells ex vivo, and  
 XX administering the tumouricidally activated T cells to the host), creating  
 XX (M5) cancer in a mammal (by administering a lipid binding molecule which  
 XX binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
 XX a tumouricidal T cell population in vivo in a mammal (by allowing a  
 XX tumour associated antigen to contact immunocytes in which adaptor  
 XX proteins which inhibit T cell activation by tumour associated antigens

CC are deleted or functionally deactivated) and producing (M7) a  
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a  
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
 CC receptors, methods and compositions are useful for treating cancers and  
 CC tumours. Bacterial superantigens are co-administered or administered as  
 CC fusion constructs with anti-tumour proteins or motifs. The present  
 CC sequence represents an anti-tumour protein which is co-administered with  
 CC or incorporated into a fusion construct with a superantigen. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format from the US patent  
 CC office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"  
 XX SQ Sequence 2796 AA;  
 Query Match 88.2%; Score 30; DB 6; Length 2796;  
 Best Local Similarity 75.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVTXIRP 9  
 DB 2704 GVHTSIRP 2711  
 RESULT 12  
 ADD31274  
 ID ADD31274 standard; peptide; 8 AA.  
 XX AC ADD31274;  
 XX DT 15-JAN-2004 (first entry)  
 XX DE Angiogenesis inhibiting peptide #49.  
 XX KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT Misc-difference 3 /note= "D-form residue"  
 FT Modified-site 5 /label= OTHER  
 FT /note= "OTHER= Nme Norvalyl (Nva)"  
 FT Modified-site 8 /note= "NHCH2CH3"  
 XX US2003109455-A1.  
 XX 12-JUN-2003.  
 XX 30-OCT-2002; 2002US-00283550.  
 XX 31-OCT-2001; 2001US-0335017P.  
 XX (HAVI/) HAVIV F.  
 XX (BRAD/) BRADLEY M F.  
 XX Haviv F, Bradley MF;  
 XX WPI; 2003-843101/78.  
 XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 XX angiogenesis, treating cancer in mammal, or for prevention of other  
 XX diseases such as autoimmune diseases.  
 XX Claim 11; Page 25; 26pp; English.  
 PS

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 XX Sequence 8 AA;  
 SQ

Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 DB 1 GVXTXIRP 8  
 |||||  
 |||||

RESULT 13  
 ADD31265  
 ID ADD31265 standard; peptide; 8 AA.  
 XX  
 AC ADD31265;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Angiogenesis inhibiting peptide #40.  
 XX  
 KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT Modified-site 3  
 FT /label= OTHER  
 FT /note= "OTHER= D-form residue, alle (not defined)"  
 FT Modified-site 5  
 FT /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT Modified-site 8  
 FT /note= "NHCH2CH3"  
 XX  
 PN US2003109455-A1.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 30-OCT-2002; 2002US-00283550.  
 XX  
 PR 31-OCT-2001; 2001US-0335017P.  
 XX  
 PA (HAVI/) HAVIV F.  
 PA (BRAD/) BRADLEY M F.  
 XX  
 PI Haviv F, Bradley MF;  
 XX  
 DR WPI; 2003-843101/78.  
 XX  
 PT New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.  
 XX  
 PS Claim 11; Page 24; 26pp; English.  
 XX

CC The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
 CC antiarthritic activities. (I) or a salt of it, is used in a  
 CC pharmaceutical composition for inhibiting angiogenesis or for treating  
 CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
 CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
 CC and degenerative arthritis. This is the amino acid sequence of an  
 CC angiogenesis inhibiting peptide of the invention.  
 XX  
 XX Sequence 8 AA;  
 SQ

Query Match 85.3%; Score 29; DB 7; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 DB 1 GVXTXIRP 8  
 |||||  
 |||||

RESULT 14  
 ADD31264  
 ID ADD31264 standard; peptide; 8 AA.  
 XX  
 AC ADD31264;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Angiogenesis inhibiting peptide #39.  
 XX  
 KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
 KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
 KW antiarthritic activities; angiogenesis inhibiting; cancer;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis.  
 XX  
 OS Synthetic.  
 XX

Key Location/Qualifiers  
 FT Modified-site 1 /label= OTHER  
 FT /note= "OTHER= N-acetyl"  
 FT Misc-difference 3  
 FT /note= "D-form residue"  
 FT Modified-site 5  
 FT /label= OTHER  
 FT /note= "OTHER= Norvalyl (Nva)"  
 FT Modified-site 8  
 FT /note= "NHCH2CH3"  
 XX  
 PN US2003109455-A1.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PF 30-OCT-2002; 2002US-00283550.  
 XX  
 PR 31-OCT-2001; 2001US-0335017P.  
 XX  
 PA (HAVI/) HAVIV F.  
 PA (BRAD/) BRADLEY M F.  
 XX  
 PI Haviv F, Bradley MF;  
 XX  
 DR WPI; 2003-843101/78.  
 XX  
 PT New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
 PT angiogenesis, treating cancer in mammal, or for prevention of other  
 PT diseases such as autoimmune diseases.  
 XX  
 PS Claim 11; Page 24; 26pp; English.  
 XX  
 CC The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
 CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and

CC antiarthritic activities. (I) or a salt of it, is used in a  
CC pharmaceutical composition for inhibiting angiogenesis or for treating  
CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
CC and degenerative arthritis. This is the amino acid sequence of an  
CC angiogenesis inhibiting peptide of the invention.  
XX  
SQ Sequence 8 AA;

Query Match 85.3%; Score 29; DB 7; Length 8;  
Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
|||  
Db 1 GVITXIRP 8

RESULT 15  
ADD31266  
ID ADD31266 standard; peptide; 8 AA.  
XX  
AC ADD31266;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE  
XX  
XX Angiogenesis inhibiting peptide #41.  
XX heptapeptide compound; octapeptide compound; nonapeptide compound;  
XX antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
XX antiarthritic activities; angiogenesis inhibiting; cancer;  
XX autoimmune disease; rheumatoid arthritis; immune arthritis;  
XX degenerative arthritis.  
XX  
OS Synthetic.

Key Location/Qualifiers  
PH Modified-site 1 /label= OTHER  
FT /note= "OTHER= N-acetyl"  
FT Misc-difference 3 /note= "D-form residue"  
FT Modified-site 4 /label= OTHER=  
FT /note= "OTHER= alloThr (not defined)"  
FT Modified-site 5 /label= OTHER  
FT /note= "OTHER= Norvalyl (Nva)"  
FT Modified-site 8 /note= "NHCH2CH3"  
XX  
XX US2003109455-A1.  
XX  
XX 12-JUN-2003.  
XX  
XX 30-OCT-2002; 2002US-00283550.  
XX  
XX 31-OCT-2001; 2001US-0335017P.  
XX  
XX (HAVI/) HAVIV F.  
XX (BRAD/) BRADLEY M F.  
XX Haviv F, Bradley MF;  
XX WPI; 2003-843101/78.  
XX  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX  
XX Claim 11; Page 24; 26pp; English.  
XX  
XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)

CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
CC antiarthritic activities. (I) or a salt of it, is used in a  
CC pharmaceutical composition for inhibiting angiogenesis or for treating  
CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
CC and degenerative arthritis. This is the amino acid sequence of an  
CC angiogenesis inhibiting peptide of the invention.  
XX  
SQ Sequence 8 AA;

Query Match 85.3%; Score 29; DB 7; Length 8;  
Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
|||  
Db 1 GVITXIRP 8

Search completed: April 1, 2004, 17:35:24  
Job time : 43.1053 secs

Sequence 11, Appli  
Sequence 5, Appli  
Sequence 6833, Ap  
Sequence 4394, A  
Sequence 25886, A  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 5475, Ap  
Sequence 13846, A  
Sequence 8, Appli  
Sequence 30148, A  
Sequence 13, Appli  
Sequence 2, Appli  
Sequence 8187, Ap  
Sequence 3, Appli  
Sequence 28762, A  
Sequence 7409, Ap

28 25 73.5 1048 4 US-09-921-099A-11  
29 24 70.6 40 2 US-08-919-724-5  
30 24 70.6 60 4 US-09-107-532A-6833  
31 24 70.6 62 4 US-09-621-976-4394  
32 24 70.6 76 4 US-09-252-991A-25888  
33 24 70.6 95 3 US-09-024-023-4  
34 24 70.6 95 4 US-09-531-111-4  
35 24 70.6 100 4 US-09-543-681A-5475  
36 24 70.6 146 4 US-09-489-039A-13846  
37 24 70.6 161 4 US-09-230-196-8  
38 24 70.6 220 4 US-09-452-991A-30148  
39 24 70.6 229 3 US-09-011-151-13  
40 24 70.6 229 4 US-09-024-023-2  
41 24 70.6 229 4 US-09-531-111-2  
42 24 70.6 252 4 US-09-328-352-8187  
43 24 70.6 269 2 US-08-727-311-3  
44 24 70.6 308 4 US-09-252-991A-28762  
45 24 70.6 323 4 US-09-489-039A-7409

ALIGNMENTS

RESULT 1  
US-08-114-555A-6  
; Sequence 6, Application US/08114555A  
; Patent No. 5854392  
; GENERAL INFORMATION:  
; APPLICANT: Manly, Susan P.  
; APPLICANT: Kozlowski, Michael R.  
; APPLICANT: Neve, Rachael L.  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100  
; TITLE OF INVENTION: RECEPTOR (C100-R)  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/114,555A  
; FILING DATE: 30-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 6013-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-114-555A-6

Query Match 85.3%; Score 29; DB 2; Length 240;  
Best Local Similarity 62.5%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GYXTXIRP 9  
||| :||

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: April 1, 2004, 17:30:38 ; Search time 12.7895 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-5  
Perfect score: 34  
Sequence: 1 XGVXTXIRP 9  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	29	85.3	240	2	US-08-114-555A-6
2	29	85.3	240	3	US-08-114-555A-8
3	29	85.3	240	3	US-08-559-397A-12
4	29	85.3	240	3	US-08-559-397A-14
5	29	85.3	748	4	US-09-252-991A-21696
6	28	82.4	271	4	US-09-252-991A-26235
7	27	79.4	1208	4	US-09-540-236-3604
8	26	76.5	216	4	US-09-252-991A-23468
9	26	76.5	479	4	US-09-252-991A-23144
10	26	76.5	728	4	US-09-252-991A-31891
11	25	73.5	208	4	US-09-252-991A-23734
12	25	73.5	293	3	US-09-307-621-4
13	25	73.5	307	4	US-09-252-991A-19676
14	25	73.5	449	1	US-08-476-008-5
15	25	73.5	449	1	US-08-476-008-7
16	25	73.5	449	1	US-08-306-063-5
17	25	73.5	449	1	US-08-306-063-7
18	25	73.5	449	1	US-08-833-485-5
19	25	73.5	449	1	US-08-833-485-7
20	25	73.5	449	3	US-09-137-440-5
21	25	73.5	449	3	US-09-137-440-7
22	25	73.5	449	5	PCT-US91-06148A-5
23	25	73.5	449	5	PCT-US91-06148A-7
24	25	73.5	596	4	US-09-252-991A-18875
25	25	73.5	836	4	US-09-252-991A-26065
26	25	73.5	929	4	US-09-252-991A-19203
27	25	73.5	947	4	US-09-719-190-2

```
Db          3  GVTSLRP 10

RESULT 2
US-08-114-555A-8
; Sequence 8, Application US/08114555A
; Patent No. 5854392
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF BETA-APP-C100
; TITLE OF INVENTION: RECEPTOR (C100-R)
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/114,555A
; FILING DATE: 30-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-114-555A-8

Query Match      85.3%; Score 29; DB 2; Length 240;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY          2  GVXTXIRP 9
Db          3  GVTSLRP 10

RESULT 3
US-08-559-397A-12
; Sequence 12, Application US/08559397A
; Patent No. 6083713
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF
; TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA

Query Match      85.3%; Score 29; DB 3; Length 240;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY          2  GVXTXIRP 9
Db          3  GVTSLRP 10

RESULT 4
US-08-559-397A-14
; Sequence 14, Application US/08559397A
; Patent No. 6083713
; GENERAL INFORMATION:
; APPLICANT: Manly, Susan P.
; APPLICANT: Kozlowski, Michael R.
; APPLICANT: Neve, Rachael L.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF
; TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA

Query Match      85.3%; Score 29; DB 3; Length 240;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY          2  GVXTXIRP 9
Db          3  GVTSLRP 10

US-08-559-397A-12
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,397A
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-559-397A-12

US-08-559-397A-14
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,397A
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6013-135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
```

; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-559-397A-14

Query Match 85.3%; Score 29; DB 3; Length 240;  
Best Local Similarity 62.5%; Pred. No. 12;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||:|  
Db 3 GVTTSLRP 10

RESULT 5  
US-09-252-991A-21696  
; Sequence 21696, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21696  
; LENGTH: 748  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-21696

Query Match 85.3%; Score 29; DB 4; Length 748;  
Best Local Similarity 75.0%; Pred. No. 39;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||:|  
Db 740 GVWTHIRP 747

RESULT 6  
US-09-252-991A-26235  
; Sequence 26235, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26235  
; LENGTH: 271  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26235

Query Match 82.4%; Score 28; DB 4; Length 271;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||:|  
Db 85 GVGTSLRP 92

RESULT 7  
US-09-540-236-3604  
; Sequence 3604, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2008-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 3604  
; LENGTH: 1208  
; TYPE: PRT  
; ORGANISM: M. catarrhalis  
US-09-540-236-3604

Query Match 79.4%; Score 27; DB 4; Length 1208;  
Best Local Similarity 50.0%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||:|  
Db 1113 GIITQVRP 1120

RESULT 8  
US-09-252-991A-23468  
; Sequence 23468, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23468  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23468

Query Match 76.5%; Score 26; DB 4; Length 216;  
Best Local Similarity 50.0%; Pred. No. 60;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||:|  
Db 69 GNATLRP 76

RESULT 9  
US-09-252-991A-23144  
; Sequence 23144, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23144
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23144

Query Match      76.5%; Score 26; DB 4; Length 479;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      219 GVSSAVRP 226

RESULT 10
US-09-252-991A-31891
; Sequence 31891, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31891
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31891

Query Match      76.5%; Score 26; DB 4; Length 728;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      474 GVSAVRP 481

RESULT 11
US-09-252-991A-23734
; Sequence 23734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23734
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23734
```

```
Query Match      73.5%; Score 25; DB 4; Length 208;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      5 GVSAVRP 12

RESULT 12
US-09-307-621-4
; Sequence 4, Application US/09307621
; Patent No. 6868487
; GENERAL INFORMATION:
; APPLICANT: Xiao, Jian-ping
; APPLICANT: Xu, Shuang-yong
; TITLE OF INVENTION: Method For Cloning And Expression of BsrFI Restriction
; TITLE OF INVENTION: Endonuclease In E. coli
; FILE REFERENCE: BsrFI
; CURRENT APPLICATION NUMBER: US/09/307,621
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Bacillus stearothermophilus
US-09-307-621-4

Query Match      73.5%; Score 25; DB 3; Length 293;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      196 GLKTSLRP 203
```

```
RESULT 13
US-09-252-991A-19676
; Sequence 19676, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19676
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19676
```

```
Query Match      73.5%; Score 25; DB 4; Length 307;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
      |||:|
Db      23 GVGTTARP 30

RESULT 14
US-08-476-008-5
; Sequence 5, Application US/08476008
```



Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyophosphate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-008-5  
Query Match 73.5%; Score 25; DB 1; Length 449;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
Db 189 GVTIVIEP 196  
RESULT 15  
US-08-476-008-7  
Sequence 7, Application US/08476008  
Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyophosphate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F

STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 449 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-008-7  
Query Match 73.5%; Score 25; DB 1; Length 449;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 GVXTXIRP 9  
Db 189 GVTIVIEP 196  
Search completed: April 1, 2004, 17:42:12  
Job time : 12.7895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:35:28 ; Search time 30.6316 Seconds  
(without alignments)  
77.074 Million cell updates/sec

Title: US-09-833-196-5  
Perfect score: 34  
Sequence: 1 XGVXTRIP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1069545 seqs, 262320428 residues

Total number of hits satisfying chosen parameters: 1069545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	88.2	89	12	US-10-424-599-188130
2	30	88.2	139	14	US-10-187-267A-17
3	30	88.2	2796	9	US-09-870-759-114
4	30	88.2	2796	10	US-09-751-708A-114
5	29	85.3	204	14	US-10-029-386-32508
6	29	85.3	312	14	US-10-156-761-11045
7	29	85.3	314	9	US-09-738-626-6809
8	29	85.3	386	12	US-10-424-599-269274
9	29	85.3	5081	12	US-10-276-774-1850
10	28	82.4	90	9	US-09-864-761-45189
11	28	82.4	147	9	US-09-864-761-46558
12	28	82.4	236	14	US-10-029-386-32561
13	28	82.4	309	15	US-10-264-237-2274
14	28	82.4	332	12	US-10-425-114-46836
15	28	82.4	376	14	US-10-156-761-8348

Sequence 5372, Ap  
Sequence 6787, Ap  
Sequence 126, App  
Sequence 2, Appli  
Sequence 4322, Ap  
Sequence 6, Appli  
Sequence 9, Appli  
Sequence 16984, A  
Sequence 280, App  
Sequence 292, App  
Sequence 9282, Ap  
Sequence 9524, Ap  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 2289, Ap  
Sequence 37683, A  
Sequence 37685, A  
Sequence 16292, A  
Sequence 83, Appli  
Sequence 15552, A  
Sequence 15925, A  
Sequence 19712, A  
Sequence 17562, A  
Sequence 1902, Ap  
Sequence 275517,  
Sequence 227151,  
Sequence 510, App  
Sequence 510, App  
Sequence 510, App  
Sequence 1422, Ap

# ALIGNMENTS

## RESULT 1

US-10-424-599-188130  
; Sequence 188130, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalick David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188130  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_140896C.1.pep  
US-10-424-599-188130

Query Match 88.2%; Score 30; DB 12; Length 89;  
Best Local Similarity 75.0%; Pred. No. 8.2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 GVXTRIP 9  
DB 78 GVITSIRP 85  
RESULT 2  
US-10-187-267A-17  
; Sequence 17, Application US/10187267A  
; Publication No. US20030124679A1  
; GENERAL INFORMATION:  
; APPLICANT: Short, Jay M.

; APPLICANT: Paraskar, Ashish  
 ; APPLICANT: Varoslu, Mustafa  
 ; APPLICANT: Mathur, Eric J.  
 ; TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF  
 ; FILE REFERENCE: 09010-280001  
 ; CURRENT APPLICATION NUMBER: US/10/187,267A  
 ; CURRENT FILING DATE: 2003-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/301,401  
 ; PRIOR FILING DATE: 2001-06-27  
 ; NUMBER OF SEQ ID NOS: 55  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 17  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces murayamaensis ATCC 21414  
 US-10-187-267A-17  
  
 Query Match 88.2%; Score 30; DB 14; Length 139;  
 Best Local Similarity 75.0%; Pred. No. 13;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 GVXTXIRP 9  
 Db 28 GVLTTIRP 35  
  
 RESULT 3  
 US-09-870-759-114  
 ; Sequence 114, Application US/09870759  
 ; Patent No. US20020177551A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TERMAN, David S  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
 ; FILE REFERENCE: 870759  
 ; CURRENT APPLICATION NUMBER: US/09/870,759  
 ; CURRENT FILING DATE: 2002-01-14  
 ; PRIOR APPLICATION NUMBER: US 60/208,128  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 166  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 114  
 ; LENGTH: 2796  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium bovis  
 US-09-870-759-114  
  
 Query Match 88.2%; Score 30; DB 9; Length 2796;  
 Best Local Similarity 75.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 GVXTXIRP 9  
 Db 2704 GVHTSIRP 2711  
  
 RESULT 4  
 US-09-751-708A-114  
 ; Sequence 114, Application US/09751708A  
 ; Publication No. US20030157113A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: TERMAN, David S  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
 ; FILE REFERENCE: 751708  
 ; CURRENT APPLICATION NUMBER: US/09/751,708A  
 ; CURRENT FILING DATE: 2002-10-15  
 ; PRIOR APPLICATION NUMBER: US 60/173,371  
 ; PRIOR FILING DATE: 1999-12-28  
 ; NUMBER OF SEQ ID NOS: 166  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 114  
 ; LENGTH: 2796  
 ; TYPE: PRT

; ORGANISM: Mycobacterium bovis  
 US-09-751-708A-114  
  
 Query Match 88.2%; Score 30; DB 10; Length 2796;  
 Best Local Similarity 75.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 GVXTXIRP 9  
 Db 2704 GVHTSIRP 2711  
  
 RESULT 5  
 US-10-029-386-32508  
 ; Sequence 32508, Application US/10029386  
 ; Publication No. US20030194704A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: AECOMICA-X-2  
 ; CURRENT APPLICATION NUMBER: US/10/029,386  
 ; CURRENT FILING DATE: 2001-12-20  
 ; NUMBER OF SEQ ID NOS: 34288  
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
 ; SEQ ID NO 32508  
 ; LENGTH: 204  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: MAP TO AC011469.3  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
 ; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.95  
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5  
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
 ; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4  
 ; OTHER INFORMATION: SWISSPROT HIT: P21817, EVALUE 1.00e-104  
 US-10-029-386-32508  
  
 Query Match 85.3%; Score 29; DB 14; Length 204;  
 Best Local Similarity 62.5%; Pred. No. 35;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
 QY 2 GVXTXIRP 9  
 Db 121 GVTTSLRP 128  
  
 RESULT 6  
 US-10-156-761-11045  
 ; Sequence 11045, Application US/10156761  
 ; Publication No. US20030119018A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OMURA, SATOSHI  
 ; APPLICANT: IKEDA, HARUO  
 ; APPLICANT: ISHIKAWA, JUN  
 ; APPLICANT: HORIKAWA, HIROSHI  
 ; APPLICANT: SHIBA, TADAYOSHI  
 ; APPLICANT: SAKAKI, YOSHIYUKI  
 ; APPLICANT: HATTORI, MASAHIRA  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-262  
 ; CURRENT APPLICATION NUMBER: US/10/156,761  
 ; CURRENT FILING DATE: 2002-05-29  
 ; PRIOR APPLICATION NUMBER: JP 2001-204089  
 ; PRIOR FILING DATE: 2001-05-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-272697  
 ; PRIOR FILING DATE: 2001-08-02  
 ; NUMBER OF SEQ ID NOS: 15109  
 ; SEQ ID NO 11045

```

; LENGTH: 312
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11045

Query Match      85.3%; Score 29; DB 14; Length 312;
Best Local Similarity 62.5%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      71 GVATVRP 78

RESULT 7
US-09-738-626-6809
; Sequence 6809, Application US/09738626
; Publication No. US20020137605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 6809
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-6809

Query Match      85.3%; Score 29; DB 9; Length 314;
Best Local Similarity 62.5%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      46 GITTFIRP 53

RESULT 8
US-10-424-599-269274
; Sequence 269274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 269274
; LENGTH: 386

```

```

; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)...(386)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85175C.1.pep
US-10-424-599-269274

Query Match      85.3%; Score 29; DB 12; Length 386;
Best Local Similarity 62.5%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      185 GVATVRP 192

RESULT 9
US-10-276-774-1850
; Sequence 1850, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 1850
; LENGTH: 5081
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-1850

Query Match      85.3%; Score 29; DB 12; Length 5081;
Best Local Similarity 62.5%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      2 GVXTXIRP 9
DB      1809 GVTTSLRP 1816

RESULT 10
US-09-864-761-45189
; Sequence 45189, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24363.6
; PRIOR FILING DATE: 2000-10-04

```

; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 45189  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL049870.3  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.46  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.52  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.54  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.48  
; OTHER INFORMATION: SWISSPROT HIT: Q04652, EVALUE 2.00e-11  
; OTHER INFORMATION: EST\_HUMAN HIT: N44974.1, EVALUE 7.00e-33  
US-09-864-761-45189

Query Match 82.4%; Score 28; DB 9; Length 90;  
Best Local Similarity 50.0%; Pred. No. 26;  
Matches 4; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 GVXTXIRP 9  
| : : :  
Db 16 GIATNVRP 23

RESULT 11  
US-09-864-761-46558  
; Sequence 46558, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wenheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 46558  
; LENGTH: 147  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL049870.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.6  
; OTHER INFORMATION: EST\_HUMAN HIT: N44974.1, EVALUE 3.00e-55  
; OTHER INFORMATION: SWISSPROT HIT: Q04652, EVALUE 1.00e-21  
US-09-864-761-46558

Query Match 82.4%; Score 28; DB 9; Length 147;  
Best Local Similarity 50.0%; Pred. No. 44;  
Matches 4; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 2 GVXTXIRP 9  
| : : :  
Db 38 GIATNVRP 45

RESULT 12  
US-10-029-386-32561  
; Sequence 32561, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32561  
; LENGTH: 236  
; TYPE: PRT  
; ORGANISM: Homo sapiens

; FEATURE:  
; OTHER INFORMATION: MAP TO AF099810.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6  
; OTHER INFORMATION: SWISSPROT HIT: P31696, EVALUE 2.00e-17  
US-10-029-386-32561

Query Match 82.4%; Score 28; DB 14; Length 236;  
Best Local Similarity 75.0%; Pred. No. 72;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
| : | : |  
Db 151 GVFTDIRP 158

RESULT 13  
US-10-264-237-2274  
; Sequence 2274, Application US/10264237  
; Publication No. US2004009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Birse et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 2274  
; LENGTH: 309  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-264-237-2274

Query Match 82.4%; Score 28; DB 15; Length 309;  
Best Local Similarity 50.0%; Pred. No. 96;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
| : | : |  
Db 76 GIATNVRP 83

RESULT 14  
US-10-425-114-46836  
; Sequence 46836, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 46836  
; LENGTH: 332  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLN017030E08\_FLI.pep  
US-10-425-114-46836

Query Match 82.4%; Score 28; DB 12; Length 332;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
| : | : |  
Db 307 GILTSVRP 314

RESULT 15  
US-10-156-761-8348  
; Sequence 8348, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 8348  
; LENGTH: 376  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-8348

Query Match 82.4%; Score 28; DB 14; Length 376;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9  
| : | : |  
Db 181 GWTVEVRP 188

Search completed: April 1, 2004, 17:45:39  
Job time : 30.6316 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 9.78947 Seconds  
(without alignments)  
88.434 Million cell updates/sec

Title: US-09-833-196-5  
Perfect score: 34  
Sequence: 1 XGVXIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	91.2	377	C82876	oligopeptide trans
2	30	88.2	2796	JC4743	fatty-acid synthas
3	29	85.3	5032	A35041	ryanodine receptor
4	29	85.3	5035	146646	ryanodine receptor
5	29	85.3	5037	E35041	ryanodine receptor
6	28	82.4	53	H84193	hypothetical prote
7	28	82.4	166	A69820	hypothetical prote
8	28	82.4	215	T44861	probable hydroxyla
9	28	82.4	368	B99563	hypothetical prote
10	28	82.4	401	A69713	septal wall dissol
11	28	82.4	420	T16919	hypothetical prote
12	28	82.4	498	H83733	hypothetical prote
13	27	79.4	141	S05804	hypothetical prote
14	27	79.4	355	D97339	dioxygenase relate
15	27	79.4	496	G87546	acid-CoA ligase, p
16	26	76.5	127	G84653	hypothetical prote
17	26	76.5	277	A39964	photosystem II oxy
18	26	76.5	288	A11966	permease protein o
19	26	76.5	313	F86937	probable pantoate-
20	26	76.5	327	I82663	transcription fact
21	26	76.5	340	T67041	hypothetical prote
22	26	76.5	359	T36781	probable gluconat
23	26	76.5	391	T36739	hypothetical prote
24	26	76.5	400	T35334	probable membrane
25	26	76.5	415	I87566	hypothetical prote
26	26	76.5	433	S05654	ND3 intron protein
27	26	76.5	520	S36489	L2 protein - human
28	26	76.5	549	T02790	hypothetical prote
29	26	76.5	597	B71130	probable oligopept

30 26 76.5 868 2 A82722 bifunctional diam  
31 26 76.5 1048 2 C86189 protein T25N20.11  
32 26 76.5 1321 2 A60165 sodium channel pro  
33 26 76.5 4967 2 S72269 ryanodine receptor  
34 26 76.5 4969 2 A37113 ryanodine receptor  
35 25 73.5 77 2 AD0343 hypothetical prote  
36 25 73.5 103 2 D71324 hypothetical prote  
37 25 73.5 135 2 T03490 conserved hypothet  
38 25 73.5 287 2 C81066 conserved hypothet  
39 25 73.5 295 1 S26978 succinate dehydrog  
40 25 73.5 296 2 AD3561 citrlyl-CoA lyase  
41 25 73.5 300 2 A81804 conserved hypothet  
42 25 73.5 314 2 C75003 na+/ca+ exchanging  
43 25 73.5 325 2 D71159 probable Na+/Ca2+-  
44 25 73.5 344 2 T19367 cysteine synthase  
45 25 73.5 347 2 A95341 probable hydrolase

## ALIGNMENTS

### RESULT 1

C82876  
oligopeptide transport system permease protein UUS62 [imported] - Ureaplasma urealyticu  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82876  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
A:Accession: C82876  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-377 <GLA>  
A:Cross-references: GB:A5002154; GB:AF222894; NID:G6899557; PIDN:AAF30975.1; GSPDB:GNOC  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: oppB; UUS62  
A:Genetic code: SGC3

Query Match 91.2%; Score 31; DB 2; Length 377;  
Best Local Similarity 75.0%; Pred. No. 3.3;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXIRP 9.  
|||  
Db 146 GVSTSRP 153

### RESULT 2

JC4743  
fatty-acid synthase (EC 2.3.1.85) - Mycobacterium bovis  
C:Species: Mycobacterium bovis  
C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 05-May-2000  
C:Accession: JC4743  
R:Fernandes, N.D.; Kolattukudy, P.E.  
Gene 170 95-99 1996  
A:Title: Cloning, sequencing and characterization of a fatty acid synthase-encoding gen  
A:Reference number: JC4743; MUID:96200863; PMID:8621098  
A:Accession: JC4743  
A:Molecule type: DNA  
A:Residues: 1-2796 <FBR>  
A:Cross-references: GB:U36763; NID:g1036834; PIDN:AA03809.1; PID:g1036835  
A:Note: the source is designated as Mycobacterium tuberculosis var. bovis BCG  
C:Comment: This enzyme catalyzes both de novo synthesis and chain elongation of fatty a  
C:Genetics:  
A:Gene: fas  
C:Superfamily: Mycobacterium tuberculosis fatty-acid synthase  
C:Keywords: acyltransferase; coenzyme A; phosphopantetheine; phosphoprotein  
F:2188-2193/Region: nucleotide binding #status predicted  
F:57/Active site: Ser #status predicted  
F:1693/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

F:2598/Active site: Cys #status predicted

Query Match 88.2%; Score 30; DB 2; Length 2796;

Best Local Similarity 75.0%; Pred. No. 52;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9

Db 2704 GVHTSLRP 2711

# RESULT 3

A35041 ryanodine receptor type 1, skeletal muscle - human

N;Alternate names: calcium release channel protein

C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jan-2000

C;Accession: A35041; I84622; S66630

R;Zorazato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meissner, G.;

J. Biol. Chem. 265, 2244-2256, 1990

A;Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release

A;Reference number: A35041; MUID:90130482; PMID:2298749

A;Accession: A35041

A;Status: Preliminary; not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-5032 <ZOR>

A;Cross-references: GB:J05200; NID:G337721; PIDN:AAA60294.1; PID:G337722

R;Otsu, K.; Phillips, M.S.; Khanna, V.K.; de Leon, S.; MacLennan, D.H.

Genomics 13, 835-837, 1992

A;Title: Refinement of diagnostic assays for a probable causal mutation for porcine and

A;Reference number: I46644; MUID:92347887; PMID:1639409

A;Accession: I46622

A;Status: translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 598-722 <RES>

A;Cross-references: GB:M91455; NID:G337723; PIDN:AAA60295.1; PID:G553643

R;Lynn, S.; Morgan, J.M.; Lamb, H.K.; Meissner, G.; Gillespie, J.I.

FEBS Lett. 372, 6-12, 1995

A;Title: Isolation and partial cloning of ryanodine-sensitive Ca(2+) release channel pro

A;Reference number: S66630; MUID:96032536; PMID:7556644

A;Accession: S66630

A;Molecule type: mRNA

A;Residues: 4690-4968 <LYN>

A;Experimental source: myometrial smooth muscle

C;Genetics:

A;Gene: GDB:RYR1

A;Cross-references: GDB:120359; OMIM:180901

A;Map position: 19q13.1-19q13.1

A;Introns: 642/2

C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

C;Keywords: calcium channel; homotetramer; phosphoprotein; skeletal muscle; transmembran

F:1788-1984/Domain: transcription initiation factor sigma region 1 homology <SR1>

Query Match 85.3%; Score 29; DB 1; Length 5032;

Best Local Similarity 62.5%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9

Db 1765 GVHTSLRP 1772

# RESULT 4

I46646

ryanodine receptor, skeletal muscle - pig

N;Alternate names: calcium release channel protein

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C;Accession: I46646; S31395; I47133; S26624; A37105; I47212; S18135

R;Fujii, J.; Otsu, K.; Zorazato, F.; de Leon, S.; Khanna, V.K.; Weiler, J.E.; O'Brien, P.

Science 253, 448-451, 1991

A;Title: Identification of a mutation in porcine ryanodine receptor associated with mal

A;Reference number: I46645; MUID:91220118; PMID:1862346

A;Accession: I46646

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-5035 <FUJ>

A;Cross-references: GB:M91452; NID:G164647; PIDN:AAA31119.1; PID:G164648

A;Accession: I46645

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-614 'C', 616-5035 <FU2>

A;Cross-references: GB:M91451; NID:G164645; PIDN:AAA31118.1; PID:G164646

R;Leeb, T.; Brem, G.; Brenig, B.

submitted to the EMBL Data Library, November 1992

A;Description: Genomic organization of porcine skeletal muscle ryanodine receptor gene

A;Reference number: S31395

A;Accession: S31395

A;Molecule type: DNA

A;Residues: 1542-2643 <LEE>

A;Cross-references: EMBL:X69465

R;Leeb, T.; Schmoelzl, S.; Brem, G.; Brenig, B.

Genomics 18, 349-354, 1993

A;Title: Genomic organization of the porcine skeletal muscle ryanodine receptor (RYR1)

A;Reference number: A48915; MUID:94117003; PMID:8288238

A;Contents: annotation

R;Harbitz, I.; Kristensen, T.; Bosnes, M.; Kran, S.; Davies, W.

Anim. Genet. 23, 395-402, 1992

A;Title: DNA sequence of the skeletal muscle calcium release channel cDNA and verificat

A;Reference number: I47133; MUID:93036581; PMID:1329581

A;Accession: I47133

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 2-2091, 'A', 2093-3494, 'L', 3496-4163, 'S', 4165-4182, 'R', 4184-4411, 'W', 4413-497

A;Cross-references: EMBL:X62880; NID:G1936; PIDN:CAA44674.1; PID:G1937

R;Harbitz, I.; Kristensen, T.; Kran, S.; Davies, W.

submitted to the EMBL Data Library, August 1992

A;Reference number: S26624

A;Accession: S26624

A;Molecule type: DNA

A;Residues: 482-706 <HAW>

A;Cross-references: EMBL:X68247

R;Harbitz, I.; Chowdhary, B.; Thomsen, P.D.; Davies, W.; Kaufmann, U.; Kran, S.; Gustav

Genomics 8, 243-248, 1990

A;Title: Assignment of the porcine calcium release channel gene, a candidate for the m

A;Reference number: A37105; MUID:91065640; PMID:2174405

A;Accession: A37105

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 4785-4971, 'R', 4973-5035 <HA2>

A;Cross-references: GB:M32501; NID:G164428; PIDN:AAA31022.1; PID:G164429

R;Ledbetter, M.W.; Preiner, J.K.; Louis, C.F.; Mickelson, J.R.

J. Biol. Chem. 269, 31544-31551, 1994

A;Title: Tissue distribution of ryanodine receptor isoforms and alleles determined by r

A;Reference number: A55660; MUID:95081095; PMID:7989322

A;Accession: I47212

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: mRNA

A;Residues: 516-676 <LED>

A;Cross-references: EMBL:U15965; NID:G52095; PIDN:AAA60467.1; PID:G52096

C;Genetics:

A;Gene: RYR1

A;Introns: 527/1; 559/3; 643/2; 1570/3; 1646/2; 1850/3; 1939/3; 2006/3; 2044/1;

A;Note: the list of introns may be incomplete

C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

C;Keywords: calcium channel

Query Match 85.3%; Score 29; DB 1; Length 5035;

Best Local Similarity 62.5%; Pred. No. 1.8e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GVXTXIRP 9

Db 1767 GVHTSLRP 1774



RESULT 5  
B35041  
ryanodine receptor, skeletal muscle - rabbit  
N/Alternate names: calcium-release channel protein; junctional channel complex  
C/Species: Oryctolagus cuniculus (domestic rabbit)  
C/Date: 27-Jul-1990 #sequence\_revision 10-Mar-1994 #text\_change 20-Aug-1999  
C/Accession: S04654; B35041; A36181; S53794; S32504  
R/Takeshima, H.; Nishimura, S.; Matsumoto, T.; Ishida, H.; Kangawa, K.; Minamino, N.; Ma  
Nature 339, 439-445, 1989  
A/Title: Primary structure and expression from complementary DNA of skeletal muscle ryan  
A/Reference number: S04654; MUID:89262082; PMID:2725677  
A/Accession: S04654  
A/Molecule type: mRNA  
A/Residues: 1-5037 <TAB>  
A/Cross-references: EMBL: X15750; NID: g1709; PIDN: CAA33762.1; PID: gi710  
A/Note: part of this sequence was confirmed by protein sequencing  
R/Zorast, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Weissner, G.;  
J. Biol. Chem. 265, 2244-2256, 1990  
A/Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release  
A/Reference number: A35041; MUID: 90130482; PMID: 2298749  
A/Accession: B35041  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-66, 'Y', 68-2014, 'D', 2016-3246, 'E', 3248-3480, 3486-4497, 'LE', 4498-4521, 'Q', 45  
4748-4758, 'N', 4760-5037 <ZOR>  
R/Marks, A.R.; Tempst, P.; Hwang, K.S.; Taubman, M.B.; Inui, M.; Chadwick, C.; Fleischer  
Proc. Natl. Acad. Sci. U.S.A. 86, 8683-8687, 1989  
A/Title: Molecular cloning and characterization of the ryanodine receptor/junctional cha  
A/Reference number: A36181; MUID: 90046857; PMID: 2813419  
A/Accession: A36181  
A/Status: nucleic acid sequence not shown; not compared with conceptual translation  
A/Molecule type: mRNA; protein  
A/Residues: 'X', 1223-1235, 'XX', 1238-1251, 1334-1348, 1566-1569, 'X', 1571, 'X', 1573, 1597-1604  
-4679, 'X', 4681-4689, 'X', 4691-4693, 'X', 4695, 'X', 4697-4700 <MAR>  
A/Note: the proteolytic fragments sequenced here from the junctional channel complex as  
R/Varsanyi, M.; Meyer, H.E.  
Biol. Chem. Hoppe-Seyler 376, 45-49, 1995  
A/Title: Sarcoplasmic reticular Ca(2+) release channel is phosphorylated at serine 2843  
A/Reference number: S53794; MUID: 95336639; PMID: 7612188  
A/Accession: S53794  
A/Molecule type: protein  
A/Residues: 2841-2852 <VAR>  
R/Takeshima, H.; Nishimura, S.; Nishi, M.; Ikeda, M.; Sugimoto, T.  
FEBS Lett. 322, 105-110, 1993  
A/Title: A brain-specific transcript from the 3'-terminal region of the skeletal muscle  
A/Reference number: S32504; MUID: 93245969; PMID: 8097730  
A/Accession: S32504  
A/Molecule type: mRNA  
A/Residues: 4163-5037 <TAW>  
C/Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog  
C/Keywords: calcium channel; glycoprotein; phosphoprotein; skeletal muscle; transmembran  
F/841-954/Region: adenine nucleotide binding  
F/1344-1359/Region: 16-residue repeats  
F/1789-1985/Domain: transcription initiation factor sigma region 1 homology <SRI>  
F/2370-2375/Region: adenine nucleotide binding  
F/2725-2844, 2845-2958/Region: 120-residue repeats  
F/4564-4580/Domain: transmembrane #status predicted <TM1>  
F/4641-4664/Domain: transmembrane #status predicted <TM2>  
F/4836-4859/Domain: transmembrane #status predicted <TM3>  
F/4918-4937/Domain: transmembrane #status predicted <TM4>  
F/128, 286/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent ki  
F/2843/Binding site: phosphate (Ser) (covalent) #status experimental  
F/4864/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.3%; Score 29; DB 2; Length 5037;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 1766 GVTSIRP 1773

RESULT 6  
H84193  
hypothetical protein Vng0346h [imported] - Halobacterium sp. NRC-1  
C/Species: Halobacterium sp. NRC-1  
C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C/Accession: H84193  
R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; I  
A/Title: Genome sequence of Halobacterium species NRC-1.  
A/Reference number: A84160; MUID: 20504483; PMID: 11016950  
A/Accession: H84193  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-53 <STO>  
A/Cross-references: GB: A8004437; NID: g10579970; PIDN: AAG18916.1; GSPDB: GN00138  
C/Genetics:  
A/Gene: VNG0346H

Query Match 82.4%; Score 28; DB 2; Length 53;  
Best Local Similarity 62.5%; Pred. No. 2.5;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 34 GVGTVRP 41

RESULT 7  
A69820  
hypothetical protein yhaw - Bacillus subtilis  
C/Species: Bacillus subtilis  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C/Accession: A69820  
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E  
Nature 390, 249-256, 1997  
A/Authors: Foulger, D.; Fritz, C.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee  
Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, J.; Sekiguchi, J.; Sero  
A/Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyana  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,  
A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A/Reference number: A69580; MUID: 98044033; PMID: 9384377  
A/Accession: A69820  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-166 <KUN>  
A/Cross-references: GB: Z99109; GB: AL009126; NID: g2633260; PIDN: CAB12823.1; PID: g2633319  
A/Experimental source: strain 168  
C/Genetics:  
A/Gene: yhaw  
C/Superfamily: Bacillus subtilis hypothetical protein yhaw

Query Match 82.4%; Score 28; DB 2; Length 166;  
Best Local Similarity 62.5%; Pred. No. 8.5;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 108 GIUTGIRP 115

```

RESULT 8
T44861
Probable hydroxylase b [imported] - Amycolatopsis orientalis (fragment)
C:Species: Amycolatopsis orientalis
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 03-Nov-2000
C:Accession: T44861
R:Solenberg, P.J.; Matsushima, P.; Stack, D.R.; Wilkie, S.C.; Thompson, R.C.; Baltz, R.H.
submitted to the EMBL Data Library, January 1997
A:Title: Production of hybrid glycopeptide antibiotics in vitro and in Streptomyces
A:Reference number: Z22861
A:Accession: T44861
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-215 <SOL>
A:Cross-references: EMBL:U84350; PIDN:AB49300.1
A:Experimental source: strain C329.4
C:Superfamily: Amycolatopsis orientalis hypothetical protein PCZA361.22

Query Match      82.4%; Score 28; DB 2; Length 215;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 2 GVXTXIRP 9
|||:|
Db 185 GVTTLLRP 192

RESULT 9
B99563
hypothetical protein MYPV 4100 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: B99563
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallisson, F.; Moszer, I.;
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: B99563
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <KUP>
A:Cross-references: GB:AL445566; PID:gl4089824; PIDN:CAC13583.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYPV 4100
A:Genetic code: SGC3

Query Match      82.4%; Score 28; DB 2; Length 368;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
|||:|
Db 137 GISTALRP 144

RESULT 10
A69713
septal wall dissolution protein spoIIP - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: A69713
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Leuber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scallan,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron

```

```

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69713
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-401 <KUN>
A:Cross-references: GB:Z99117; GB:AL009126; NID:g2634966; PIDN:CAB14495.1; PID:g2634995
A:Experimental source: strain 168
C:Genetics:
A:Gene: spoIIP

Query Match      82.4%; Score 28; DB 2; Length 401;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
|||:|
Db 38 GVLTSLRP 45

RESULT 11
T16919
hypothetical protein T21F4.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 18-Feb-2000
C:Accession: T16919
R:Minx, P.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid T21F4.
A:Reference number: Z18605
A:Accession: T16919
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-420 <MIN>
A:Cross-references: EMBL:U56959; NID:gl293801; PID:gl293802; PIDN:AAA98709.1; GSPDB:GNC
A:Experimental source: strain Bristol N2; clone T21F4
C:Genetics:
A:Gene: CBSP:T21F4.1
A:Map position: X
A:Introns: 73/3; 128/3; 185/3; 219/2; 313/3; 386/1

Query Match      82.4%; Score 28; DB 2; Length 420;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
|||:|
Db 108 GVATMRP 115

RESULT 12
H83793
hypothetical protein BH152 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H83793
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H83793
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-498 <STO>
A:Cross-references: GB:AP001511; GB:BA000004; NID:gi0173727; PIDN:BA004871.1; GSPDB:GNO
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH152

Query Match      82.4%; Score 28; DB 2; Length 498;

```

```

Best Local Similarity 62.5%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 107 GVLTVGRP 114

RESULT 13
hypothetical protein UL41 - human cytomegalovirus (strain AD169)
C:Species: human cytomegalovirus, human herpesvirus 5
A:Note: host Homo sapiens [man]
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Jul-2000
C:Accession: S09804
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsneil, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A:Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A:Reference number: S09749; MUID:90269039; PMID:2161319
A:Accession: S09804
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-141 <CHE>
A:Cross-references: EMBL:X17403; NID:959591; PIDN:CAA35400.1; PID:e27240; PID:g1813967
A:Note: this sequence was submitted to the EMBL Data Library, December 1989
A:Note: this reading frame extends between two stop codons and does not begin with a sta
C:Superfamily: human cytomegalovirus hypothetical protein UL41

Query Match 79.4%; Score 27; DB 2; Length 141;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 8 GIITLTP 15

RESULT 14
D97339
dioxxygenase related to 2-nitropropane dioxxygenase [imported] - Clostridium acetobutylicu
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: D97339
R:Nolling, J.; Berton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4836, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: D97339
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-355 <KUP>
A:Cross-references: GB:AE001437; PIDN:AAK81503.1; PID:g15026676; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
C:Superfamily: Campylobacter jejuni hypothetical protein Cj1270c

Query Match 79.4%; Score 27; DB 2; Length 355;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 190 GVLTAIRP 197

RESULT 15
GS7546
acid-CoA ligase, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

```

```

C:Accession: G87546
R:Nierman, W.C.; DeBolyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolc
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-496 <STO>
A:Cross-references: GB:AE005673; NID:g13423937; PIDN:AAK24371.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2400
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 79.4%; Score 27; DB 2; Length 496;
Best Local Similarity 62.5%; Pred. No. 51;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
Db 196 GLITSIRP 203

```

Search completed: April 1, 2004, 17:40:42  
Job time : 10.7895 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:18 ; Search time 5.52632 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-5  
Perfect score: 34  
Sequence: 1 XGVYXIRP 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	85.3	5035	1 RYR1_PIG	P16960 sus scrofa
2	29	85.3	5037	1 RYR1_RABIT	P11716 oryctolagus
3	29	85.3	5038	1 RYR1_HUMAN	P21817 homo sapien
4	28	82.4	401	1 SP2P_BACSU	P37968 bacillus su
5	28	82.4	1541	1 NX3A_HUMAN	Q9Y4C0 homo sapien
6	27	79.4	141	1 UL41_HCMVA	P16814 human cytom
7	26	76.5	277	1 PSBO_SYNP7	P11472 synchococc
8	26	76.5	343	1 PANC_MYCLE	O69524 mycobacteri
9	26	76.5	340	1 YD60_MYCTU	Q11030 mycobacteri
10	26	76.5	370	1 PIT1_MELGA	Q05749 meleagris g
11	26	76.5	375	1 ADRI_HUMAN	Q96A54 homo sapien
12	26	76.5	375	1 ADRI_MOUSE	Q91VH1 mus musculu
13	26	76.5	433	1 YMN3_PODAN	P15563 podospora a
14	26	76.5	520	1 VL2_HPV19	P36752 human papil
15	26	76.5	4967	1 RYR2_HUMAN	Q92736 homo sapien
16	26	76.5	4969	1 RYR2_RABIT	P30957 oryctolagus
17	25	73.5	295	1 DHSB_USTMA	P23420 ustilago ma
18	25	73.5	314	1 O2D3_HUMAN	Q89H33 homo sapien
19	25	73.5	449	1 AROA_PSES2	P56952 pseudomonas
20	25	73.5	450	1 AROA_BRUB	Q9AGV2 brucella ab
21	25	73.5	450	1 AROA_BRUME	O8Y9G1 brucella me
22	25	73.5	450	1 AROA_BRUSU	O8G5C4 brucella su
23	25	73.5	466	1 WR32_ARATH	P59583 arabidopsis
24	25	73.5	511	1 PUR9_SYNY3	P74741 s bifunctio
25	25	73.5	913	1 YC17_HABIN	P45114 haemophilus
26	25	73.5	923	1 AGLU_TETPY	O00906 tetrahymena
27	25	73.5	947	1 2268_HUMAN	Q14587 homo sapien
28	25	73.5	1017	1 EM11_MOUSE	O99K41 mus musculu
29	25	73.5	2355	1 TOR1_SCHPO	O14356 schizosach
30	25	73.5	2565	1 TRAP_MOUSE	Q80YV3 mus musculu
31	25	73.5	3859	1 TRAP_HUMAN	Q9Y4A5 homo sapien
32	24	70.6	134	1 IF2B_PYRAE	Q82X46 pyrobaculum
33	24	70.6	143	1 YF46_MYCTU	Q10780 mycobacteri

#### ALIGNMENTS

##### RESULT 1

ID	RYR1_PIG	STANDARD;	PRT;	5035 AA.
AC	P16960;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RYR1)			
DE	(RYR-1) (Skeletal muscle calcium release channel).			
GN	RYR1 OR CRC			
OS	Sus scrofa (Pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
OX	NCBI_TaxID=9823;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Norwegian Landrace; TISSUE=Skeletal muscle;			
RX	MEDLINE=9303658; PubMed=1329581;			
RA	Harbitz I., Kristensen T., Bosnes M., Kran S., Davies W.;			
RT	"DNA sequence of the skeletal muscle calcium release channel cDNA and			
RT	verification of the Arg615-->Cys615 mutation, associated with porcine			
RT	malignant hyperthermia, in Norwegian landrace pigs.";			
RL	Anim. Genet. 23:395-402(1992).			
RN	[2]			
RP	SEQUENCE OF 1129-2801 FROM N.A.			
RC	Brenig B.;			
RL	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 1129-2643 FROM N.A.			
RC	STRAIN=German Landrace; TISSUE=Liver;			
RX	MEDLINE=94117003; PubMed=8288238;			
RA	Leeb T., Schmolzi S., Brem G., Brenig B.;			
RT	"Genomic organization of the porcine skeletal muscle ryanodine			
RT	receptor (RYR1) gene coding region 4624 to 7929.";			
RL	Genomics 18:349-354(1993).			
RN	[4]			
RP	SEQUENCE OF 4785-5035 FROM N.A.			
RX	MEDLINE=91065640; PubMed=2174405;			
RA	Harbitz I., Chowdhary B., Thomsen P.D., Davies W., Kaufman U.,			
RA	Kran S., Gustavsen I., Christensen K., Hauge J.G.;			
RT	"Assignment of the porcine calcium release channel gene, a candidate			
RT	for the malignant hyperthermia locus, to the 6p11-->q21 segment of			
RT	chromosome 6.";			
RL	Genomics 8:243-248(1990).			
CC	-!- FUNCTION: Communication between transverse-tubules and			
CC	sarcoplasmic reticulum. Contraction of skeletal muscle is			
CC	triggered by release of calcium ions from SR following			
CC	depolarization of T-tubules.			
CC	-!- SUBUNIT: Homotrimer (Potential).			
CC	-!- MISCELLANEOUS: The calcium release channel is modulated by calcium			
CC	ions, magnesium ions, ATP and calmodulin.			
CC	-!- MISCELLANEOUS: The calcium release channel activity resides in the			
CC	C-terminal region while the remaining part of the protein			
CC	constitutes the 'foot' structure spanning the junctional gap			
CC	between the SR and the T-tubule. It is possible that the foot			
CC	structure interacts with the cytoplasmic region of the			

Q99sf2 staphylococ  
Q0463 arabidopeis  
Q89k3 buchnera ap  
Q9hz0 pseudomonas  
Q881c3 pseudomonas  
Q87V42 pseudomonas  
Q00115 ictalurid h  
Q8zeu2 yersinia pe  
P38809 saccharomyc  
Q7v0r8 prochloroco  
O59930 phanerochae  
Q09551 caenorhabdi

34 24 70.6 179 1 ATPD\_STAAM  
35 24 70.6 198 1 IFP2\_ARATH  
36 24 70.6 238 1 PLGH\_BUCAP  
37 24 70.6 295 1 PNK\_PSEAE  
38 24 70.6 296 1 PNK\_PSEPK  
39 24 70.6 296 1 PNK\_PSESM  
40 24 70.6 302 1 VG03\_HSV11  
41 24 70.6 318 1 ZNUA\_YERPE  
42 24 70.6 366 1 YHP7\_YEAST  
43 24 70.6 379 1 OXAA\_PROMP  
44 24 70.6 380 1 LEU3\_PHACH  
45 24 70.6 401 1 YQV1\_CAEEL

```

CC      dihydropyridine receptor'.
CC      -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-
CC      release channel in junctional SR and modulates its activity.
CC      -!- SIMILARITY: Belongs to the ryanodine receptor family.
CC      -!- SIMILARITY: Contains 5 MIR domains.
CC      -!- SIMILARITY: Contains 3 SPRY domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sb-sib.ch).

```

CC dihydropyridine receptor.  
 CC -1- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-  
 CC release channel in junctional SR and modulates its activity.  
 CC -1- SIMILARITY: Belongs to the ryanodine receptor family.  
 CC -1- SIMILARITY: Contains 5 MIR domains.  
 CC -1- SIMILARITY: Contains 3 SPRY domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X15209; CAA33279.1; -;  
 CC EMBL; X15750; CAA33762.1; -;  
 CC EMBL; S04654; B35041.  
 CC InterPro: IPR000699; Ca-rel channel.  
 CC InterPro: IPR001682; Ca-ha.Pore.  
 CC InterPro: IPR002048; EF-hand.  
 CC InterPro: IPR005821; Ion trans.  
 CC InterPro: IPR003608; MIR.  
 CC InterPro: IPR001215; Ryanodn\_receptor.  
 CC InterPro: IPR003032; RYR.  
 CC InterPro: IPR003877; SPRY receptor.  
 CC Pfam; PF00520; ion trans; 1.  
 CC Pfam; PF02815; MIR; 4.  
 CC Pfam; PF01365; RYR; 1TPR; 2.  
 CC Pfam; PF02026; RYR; 4.  
 CC Pfam; PF00622; SPRY; 3.  
 CC PRINTS; PR00795; RYANODINER.  
 CC SMART; SM00472; MIR; 4.  
 CC SMART; SM00449; SPRY; 3.  
 CC PROSITE; PS50919; MIR; 5.  
 CC Receptor, Transmembrane; Ionic channel; Calcium channel; Repeat;  
 KW Phosphorylation; Glycoprotein.  
 FT DOMAIN 1 3123 CYTOPLASMIC.  
 FT TRANSMEM 3124 3144 M' (POTENTIAL).  
 FT TRANSMEM 3188 3206 M'' (POTENTIAL).  
 FT TRANSMEM 3985 4004 M1 (POTENTIAL).  
 FT TRANSMEM 4023 4041 M2 (POTENTIAL).  
 FT TRANSMEM 4277 4300 M3 (POTENTIAL).  
 FT TRANSMEM 4342 4362 M4 (POTENTIAL).  
 FT TRANSMEM 4559 4580 M5 (POTENTIAL).  
 FT TRANSMEM 4648 4671 M6 (POTENTIAL).  
 FT TRANSMEM 4789 4809 M7 (POTENTIAL).  
 FT TRANSMEM 4837 4856 M8 (POTENTIAL).  
 FT TRANSMEM 4879 4898 M9 (POTENTIAL).  
 FT TRANSMEM 4914 4937 M10 (POTENTIAL).  
 FT DOMAIN 98 153 MIR 1.  
 FT DOMAIN 160 205 MIR 2.  
 FT DOMAIN 211 265 MIR 3.  
 FT DOMAIN 271 334 MIR 4.  
 FT DOMAIN 336 394 MIR 5.  
 FT DOMAIN 659 797 SPRY 1.  
 FT DOMAIN 1085 1208 SPRY 2.  
 FT DOMAIN 1430 1570 SPRY 3.  
 FT DOMAIN 1873 1913 GLU-RICH (ACIDIC).  
 FT DOMAIN 842 2959 6 X APPROXIMATE REPEATS.  
 FT REPEAT 956 955 1.  
 FT REPEAT 1345 1360 2.  
 FT REPEAT 1373 1388 3 (INCOMPLETE).  
 FT REPEAT 2726 2845 4 (INCOMPLETE).  
 FT REPEAT 2846 2959 5.  
 FT MOD RES 2843 2843 PHOSPHORYLATION (BY PKA AND PKG).  
 FT MOD RES 3952 3952 PHOSPHORYLATION (POTENTIAL).  
 FT MOD RES 4323 4323 PHOSPHORYLATION (POTENTIAL).  
 FT CARBOHYD 3466 3466 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3909 3909 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3950 3950 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 4149 4149 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 4864 4864 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 2015 E -> D (IN REF. 2).  
 FT CONFLICT 3481 3485 MISSING (IN REF. 2).  
 SQ SEQUENCE 5037 AA; 565238 MW; 4ABD87AA2697070 CRC64;  
 Query Match 85.3%; Score 29; DB 1; Length 5037;  
 Best Local Similarity 62.5%; Pred. No. 90;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVXTXIRP 9  
 DB 1766 GVTTSLRP 1773  
 RESULT 3  
 RYR1\_HUMAN  
 ID RYR1\_HUMAN STANDARD; PRT; 5038 AA.  
 AC P21817; Q16314; Q16368; Q9NP1K; Q9PLU4;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RYR1)  
 DE (RYR-1) (Skeletal muscle calcium release channel).  
 GN RYR1 OR RYDR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND PARTIAL SEQUENCE.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=90130482; PubMed=2298749;  
 RA Zorzato F., Fujii J., Otsu K., Phillips M.S., Green N.M., Lai F.A.,  
 RA Meissner G., MacLennan D.H.;  
 RT "Molecular cloning of cDNA encoding human and rabbit forms of the  
 RT Ca2+ release channel (ryanodine receptor) of skeletal muscle  
 RT sarcoplasmic reticulum";  
 RL J. Biol. Chem. 265:2244-2256(1990).  
 RN [2]  
 RP REVISIONS TO 2324; 2840 AND 3380, VARIANT MH ARG-248, AND VARIANTS  
 RP CYS-471; LEU-1787 AND CYS-2060.  
 RC TISSUE=Muscle;  
 RX MEDLINE=92372020; PubMed=1354642;  
 RA Gillard E.F., Otsu K., Fujii J., Duff C.L., de Leon S., Khanna V.K.,  
 RA Britt B.A., Worton R.G., McLennan D.H.;  
 RT "Polymorphisms and deduced amino acid substitutions in the coding  
 RT sequence of the ryanodine receptor (RYR1) gene in individuals with  
 RT malignant hyperthermia";  
 RL Genomics 13:1247-1254(1992).  
 RN [3]  
 RP REVISIONS TO 1365-1368, VARIANT CCD/MH HIS-2435, AND ALTERNATIVE  
 RP SPLICING.  
 RC TISSUE=Muscle;  
 RX MEDLINE=94035117; PubMed=8220422;  
 RA Zhang Y., Chen H.S., Khanna V.K., de Leon S., Phillips M.S.,  
 RA Schappert K., Britt B.A., Brownell A.K.W., McLennan D.H.;  
 RT "A mutation in the human ryanodine receptor gene associated with  
 RT central core disease";  
 RL Nat. Genet. 5:46-50(1993).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=96299657; PubMed=8661021;  
 RA Phillips M.S., Fujii J., Khanna V.K., de Leon S., Yokobata K.,  
 RA de Jong P.J., McLennan D.H.;  
 RT "The structural organization of the human skeletal muscle ryanodine  
 RT receptor (RYR1) gene";  
 RL Genomics 34:24-41(1996).  
 RN [5]  
 RP SEQUENCE OF 598-722 FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX MEDLINE=92347887; PubMed=1639409;  
 RA Otsu K., Phillips M.S., Khanna V.K., de Leon S., McLennan D.H.;  
 RT "Refinement of diagnostic assays for a probable causal mutation for

porcine and human malignant hyperthermia.";  
 Genomics 13:835-837(1992).  
 [6]  
 RP SEQUENCE OF 3339-5038 FROM N.A. (ISOFORM 2).  
 RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
 RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
 RA Phan H., Velasco N., Do L., Regala W., Terry A., Gaines J.,  
 RA Danganan L., Exler A., Christensen M., Georgescu A., Avila J.,  
 RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,  
 RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
 RA Krommiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,  
 RA Kobayashi A., Olsen A.S., Carrano A.V.,  
 RA Submitted (NOV-1998) to the ENBL/GenBank/DBJ databases.  
 [7]  
 RP SEQUENCE OF 4696-4974 FROM N.A.  
 RA TISSUE=Myometrium;  
 RC MEDLINE=96032536; PubMed=7556644;  
 RA Lynn S., Morgan J.M., Lamb H.K., Weisner G., Gillespie J.I.;  
 RT "Isolation and partial cloning of ryanodine-sensitive Ca2+ release  
 channel protein isoforms from human myometrial smooth muscle.";  
 RL PNAS Lett. 372:16-12(1995).  
 [8]  
 RP TISSUE SPECIFICITY.  
 RA MEDLINE=98268728; PubMed=9607712;  
 RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;  
 RT "Partial cloning and differential expression of ryanodine  
 receptor/calcium-release channel genes in human tissues including  
 the hippocampus and cerebellum.";  
 RL Neuroscience 85:205-216(1998).  
 [9]  
 RP S-NITROSYLATION OF CYS-3635.  
 RA MEDLINE=21457268; PubMed=11562475;  
 RA Sun J., Xin C., Eu J.P., Stamlor J.S., Weissner G.;  
 RT "Cysteine-3635 is responsible for skeletal muscle ryanodine receptor  
 modulation by NO.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:11158-11162(2001).  
 [10]  
 RP VARIANT MH CYS-614.  
 RA MEDLINE=92128959; PubMed=1774074;  
 RA Gillard E.F., Otsu K., Fujii J., Khanna V.K., de Leon S.,  
 RA Dardemazi J., Britt B.A., Duff C.L., Worton R.G., McLennan D.H.;  
 RT "A substitution of cysteine for arginine 614 in the ryanodine  
 receptor is potentially causative of human malignant hyperthermia.";  
 RL Genomics 11:751-755(1991).  
 [11]  
 RP VARIANTS CDS/MH CYS-163 AND MET-403.  
 RA MEDLINE=94035118; PubMed=8220423;  
 RA Quane K.A., Healy J.M.S., Keating K.E., Manning B.M., Couch F.J.,  
 RA Palmucci L.M., Doriguzzi C., Fagerlund T.H., Berg K., Ording H.,  
 RA Bendixen D., Mortier W., Linz U., Muller C.R., McCarthy T.V.;  
 RT "Mutations in the ryanodine receptor gene in central core disease and  
 malignant hyperthermia.";  
 RL Nat. Genet. 5:51-55(1993).  
 [12]  
 RP VARIANT CDS/MH SER-522.  
 RA MEDLINE=95130087; PubMed=7829078;  
 RA Quane K.A., Keating K.E., Healy J.M.S., Manning B.M.,  
 RA Krivosic-Horber R., Krivosic I., Monnier N., Lunardi J.,  
 RA McCarthy T.V.;  
 RT "Mutation screening of the RYR1 gene in malignant hyperthermia:  
 RT detection of a novel Tyr to Ser mutation in a pedigree with  
 associated central cores.";  
 RL Genomics 23:236-239(1994).  
 [13]  
 RP VARIANT MH ARG-341.  
 RA MEDLINE=94282042; PubMed=8012359;  
 RA Quane K.A., Keating K.E., Manning B.M., Healy J.M.S., Monsteurs K.,  
 RA Heffron J.J.A., Lehane M., Heytens L., Krivosic-Horber R., Adnet P.,  
 RA Ellis F.R., Monnier N., Lunardi J., McCarthy T.V.,  
 RA "Detection of a novel common mutation in the ryanodine receptor gene  
 RT in malignant hyperthermia: implications for diagnosis and  
 RT heterogeneity studies.";  
 RL Hum. Mol. Genet. 3:471-476(1994).

[14]  
 RP VARIANT MH ARG-2434.  
 RA MEDLINE=95152512; PubMed=7849712;  
 RA Keating K.E., Quane K.A., Manning B.M., Lehane M., Hartung E.,  
 RA Censier K., Urwyler A., Klausnitzer M., Muller C.R., Heffron J.J.A.,  
 RA McCarthy T.V.;  
 RT "Detection of a novel RYR1 mutation in four malignant hyperthermia  
 RT pedigrees.";  
 RL Hum. Mol. Genet. 3:1855-1858(1994).  
 [15]  
 RP VARIANT MH ARG-2434.  
 RA MEDLINE=95187158; PubMed=7881417;  
 RA Phillips M.S., Khanna V.K., de Leon S., Britt B.A.,  
 RA McLennan D.H.;  
 RT "The substitution of Arg for Gly2433 in the human skeletal muscle  
 RT ryanodine receptor is associated with malignant hyperthermia.";  
 RL Hum. Mol. Genet. 3:2181-2186(1994).  
 [16]  
 RP VARIANT MH CYS-614.  
 RA TISSUE=Blood;  
 RC MEDLINE=95271229; PubMed=7751854;  
 RA Moroni I., Gonano E.F., Comi G.P., Tegazzin V., Pirelli A., Bordini A.,  
 RA Bresolin N., Scarlato G.;  
 RT "Ryanodine receptor gene point mutation and malignant hyperthermia  
 RT susceptibility.";  
 RL J. Neurol. 242:127-133(1995).  
 [17]  
 RP VARIANT MH ARG-35.  
 RA MEDLINE=97219028; PubMed=9066328;  
 RA Lynch P.J., Krivosic-Horber R., Reyford H., Monnier N., Quane K.A.,  
 RA Adnet P., Haudecoeur G., Krivosic I., McCarthy T.V., Lunardi J.;  
 RT "Identification of heterozygous and homozygous individuals with the  
 RT novel RYR1 mutation Cys35Arg in a large kindred.";  
 RL Anesthesiology 88:620-626(1997).  
 [18]  
 RP VARIANT MH LEU-614.  
 RA MEDLINE=98051290; PubMed=9389651;  
 RA Quane K.A., Ording H., Keating K.E., Manning B.M., Heine R.,  
 RA Bendixen D., Berg K., Krivosic-Horber R., Lehmann-Horn F.,  
 RA Fagerlund T.H., McCarthy T.V.;  
 RT "Detection of a novel mutation at amino acid position 614 in the  
 RT ryanodine receptor in malignant hyperthermia.";  
 RL Br. J. Anaesth. 79:332-337(1997).  
 [19]  
 RP VARIANT MH TRP-552.  
 RA MEDLINE=97284075; PubMed=9138151;  
 RA Keating K.E., Giblin L., Lynch P.J., Quane K.A., Lehane M.,  
 RA Heffron J.J.A., McCarthy T.V.;  
 RT "Detection of a novel mutation in the ryanodine receptor gene in an  
 RT Irish malignant hyperthermia pedigree: correlation of the IVCT  
 RT response with the affected and unaffected haplotypes.";  
 RL J. Med. Genet. 34:231-236(1997).  
 [20]  
 RP VARIANTS MH CYS-2163, MET-2168 AND MET-2206, AND VARIANT CDS/MH  
 HIS-2163.  
 RA MEDLINE=98163444; PubMed=9497245;  
 RA Manning B.M., Quane K.A., Ording H., Urwyler A., Tegazzin V.,  
 RA Lehane M., O'Halloran J., Hartung E., Giblin L.M., Lynch P.J.,  
 RA Vaughan P., Censier K., Bendixen D., Comi G.P., Heytens L.,  
 RA Monsteurs K., Fagerlund T.H., Wolz W., Heffron J.J.A., Mueller C.R.,  
 RA McCarthy T.V.;  
 RT "Identification of novel mutations in the ryanodine receptor gene  
 RT (RYR1) in malignant hyperthermia: genotype-phenotype correlation.";  
 RL Am. J. Hum. Genet. 62:599-609(1998).  
 [21]  
 RP VARIANTS MH CYS-2459 AND HIS-2459.  
 RA MEDLINE=98111378; PubMed=9450902;  
 RA Manning B.M., Quane K.A., Lynch P.J., Urwyler A., Tegazzin V.,  
 RA Krivosic-Horber R., Censier K., Comi G.P., Adnet P., Wolz W.,  
 RA Lunardi J., Muller C.R., McCarthy T.V.;

Query Match 85.3%; Score 29; DB 1; Length 5038;  
 Best Local Similarity 62.5%; Pred. No. 90;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
 DB 1766 GVLTSLRP 1773

RESULT 4  
 SP2P\_BACSU STANDARD; PRT; 401 AA.  
 AC P3796;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Stage II sporulation protein P.  
 GN SPOIIP OR BSU25530.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RA Takemaru K.I., Sato T., Kobayashi Y.;  
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / JH642;  
 RX MEDLINE=97124195; PubMed=8969508;  
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,  
 RA Kobayashi Y.;  
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of  
 RT the Bacillus subtilis genome containing the skin element and many  
 RT sporulation genes";  
 RL Microbiology 142:3103-3111 (1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Besieres P., Bolotin A., Borchert S.,  
 RA Bories R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Broutlet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton J.F., Cummings N.D., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Faret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holgappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Furnelle B., Rapoport G., Rey M., Sadale Y.,  
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis";  
 RL Nature 390:249-256 (1997).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D17650; EAA04542.1; -;  
 DR EMBL; D84432; EAA12458.1; -;  
 DR EMBL; Z99117; CAB14495.1; -;  
 DR PIR; A69713; A69713;  
 DR Subtilis; BG10439; spoIIP.  
 KW Sporulation; Complete proteome.  
 SQ SEQUENCE 401 AA; 44548 MW; EB060014088E17A5 CRC64;  
 Query Match 82.4%; Score 28; DB 1; Length 401;  
 Best Local Similarity 62.5%; Pred. No. 14;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVXTXIRP 9  
 DB 38 GVLTSLRP 45

RESULT 5  
 NX3A\_HUMAN STANDARD; PRT; 1541 AA.  
 ID NX3A\_HUMAN Q9Y4C0; Q95378; Q9PLV3; Q9PLV6; Q9UIE2; Q9UIE3; Q9ULAS;  
 AC Q9Y486;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Neurexin 3-alpha precursor (Neurexin III-alpha).  
 DE NRXN3 OR KIAA0743.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21945268; PubMed=11944992;  
 RX Rowen L., Young J., Birditt B., Kaur A., Madan A., Philipps D.L.,  
 RA Qin S., Minx P., Wilson R.K., Hood L., Graveley B.R.;  
 RT "Analysis of the human neurexin genes: alternative splicing and the  
 RT generation of protein diversity";  
 RL Genomics 79:587-597 (2002).  
 RN [2]  
 RP SEQUENCE OF 253-1541 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=99087487; PubMed=9872452;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,  
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XI.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro";  
 RL DNA Res. 5:277-286 (1998).  
 CC -1- FUNCTION: Neuronal cell surface protein that may be involved in  
 CC cell recognition and cell adhesion. May mediate intracellular  
 CC signaling.  
 CC -1- SUBUNIT: The laminin G-like domain 2 binds to neuexophilin 1.  
 CC Specific isoforms bind to alpha-dystroglycan. The cytoplasmic C-  
 CC terminal region binds to CASK (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative promoter;  
 CC Comment=A number of isoforms, alpha-type (shown here) and  
 CC beta-type (AC Q9HDB5), are produced by use of alternative  
 CC promoters. Beta-type isoforms differ from alpha-type isoforms  
 CC in their N-terminus;  
 CC Event=Alternative splicing; Named isoforms=1;  
 CC Comment=A number of isoforms may be produced;  
 CC Name=1;  
 CC IsoId=Q9Y4C0-1; Sequences=Displayed;  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in brain.  
 CC -1- SIMILARITY: Contains 6 laminin G-like domains.



--!- SIMILARITY: Contains 3 EGF-like domains.  
--!- SIMILARITY: Belongs to the neurexin family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; AF099810; AAC68909.1; -.  
EMBL; AC008045; AAF28465.1; -.  
EMBL; AC009396; AAF21147.1; -.  
EMBL; AC012099; AAF15058.1; -.  
EMBL; AC026888; AAF87841.1; -.  
EMBL; AC011440; AAF61277.1; -.  
EMBL; AC008056; AAF09143.1; -.  
EMBL; AF123462; AAD13621.1; -.  
EMBL; AB018286; BAA34463.1; ALT INIT.  
HSP; Q63373.1C4R  
Genew; HGNC:8010; NRXN3.  
MIM; 600567; -.  
GO; GO:0005887; C:integral to plasma membrane; TAS.  
GO; GO:0004872; F:receptor activity; TAS.  
GO; GO:0007411; P:axon guidance; TAS.  
InterPro; IPR000152; Asx hydroxyl S.  
InterPro; IPR008985; ConA\_like\_iec\_gl.  
InterPro; IPR006209; EGF-like.  
InterPro; IPR001791; Laminin G.  
Pfam; PF00008; EGF; 3.  
Pfam; PF00054; laminin G; 6.  
PROSITE; PS00010; ASX\_HYDROXYL; 1.  
PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
PROSITE; PS01186; EGF\_2; FALSE\_NEG.  
PROSITE; PS00026; EGF\_3; 3.  
PROSITE; PS50025; LAM\_G\_DOMAIN; 6.  
Signal; Transmembrane; Repeat; Cell adhesion; EGF-like domain;  
Glycoprotein; Alternative splicing; Alternative promoter usage.  
KW SIGNAL 1 27  
FT CHAIN 28 1541  
FT DOMAIN 28 1466  
FT TRANSMEM 1467 1487  
FT DOMAIN 1488 1541  
FT DOMAIN 28 202  
FT DOMAIN 138 235  
FT DOMAIN 258 440  
FT DOMAIN 447 639  
FT DOMAIN 643 680  
FT DOMAIN 685 857  
FT DOMAIN 871 1046  
FT DOMAIN 1049 1086  
FT DOMAIN 1090 1260  
FT DOMAIN 1269 1292  
FT DOMAIN 1295 1298  
FT DOMAIN 1370 1373  
FT DOMAIN 1385 1388  
FT DOMAIN 1474 1477  
FT CARBOHYD 58 58  
FT CARBOHYD 105 105  
FT CARBOHYD 757 757  
FT CARBOHYD 1189 1189  
FT CARBOHYD 1257 1257  
FT CARBOHYD 1301 1301  
FT CARBOHYD 1043 1043  
FT CONFLICT 1334 1334

K -> E (IN REF. 1; AAF87841/AAF61277).  
T -> TTGGELVIFLLVEDPLATPPATRAPSLTPPTFRP  
LTITETTKDSMTSEAGLCFSDSGSCDDGLVISGY  
GSGETFDNLPTDDEDFYTFLSVLT (IN REF. 1;  
RAAI3621).

SEQUENCE 1541 AA; 169923 MW; BE67FE2FE6197C95 CRC64;

Query Match 82.4%; Score 28; DB 1; Length 1541.  
Best Local Similarity 75.0%; Pred No. 52;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
DB 151 GVPTDIRP 158

RESULT 6  
UL41\_HCMVA STANDARD; PRT; 141 AA.  
ID UL41\_HCMVA  
AC F168J4;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Hypothetical protein UL41.  
GN UL41.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
NCBI\_TaxID=10360;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=90269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,  
RA Reddell E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
RT Analysis of the protein-coding content of the sequence of human  
RT cytomegalovirus strain AD169.";  
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X17403; CAA35400.1; --  
DR PIR; S09804; S09804.  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 16766 MW; 20005377B1EFB712 CRC64;  
  
Query Match 79.4%; Score 27; DB 1; Length 141;  
Best Local Similarity 50.0%; Pred. No. 9.4;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GVXTXIRP 9  
DB 8 GIITLRP 15

RESULT 7  
PSBO\_SYNPF STANDARD; PRT; 277 AA.  
ID PSBO\_SYNPF  
AC P11472;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Photosystem II manganese-stabilizing polypeptide precursor (MSP).  
DE PSBO OR W0XA.  
GN Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).  
OS Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
NCBI\_TaxID=1140;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=86069567; PubMed=3120187;  
RA Kuwabara I., Reddy K.J., Sherman L.A.;  
RT "Nucleotide sequence of the gene from the cyanobacterium Anacystis  
RT nidulans R2 encoding the Mn-stabilizing protein involved in  
RT photosystem II water oxidation.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8230-8234(1987).  
[2]

```
Query Match      82.4%; Score 28; DB 1; Length 1541;
Best Local Similarity 75.0%; Pred. No. 52;
```

SEQUENCE OF 28-41.  
 RX MEDLINE=90092104; PubMed=2689172;  
 RA Kuwabara T., Nagata R., Shinohara K.;  
 RT "Expression and processing of cyanobacterial Mn-stabilizing protein  
 in *Escherichia coli*.";  
 RL Eur J Biochem. 186:227-232(1989).  
 CC -!- FUNCTION: MSP BINDS TO A PUTATIVE MN-BINDING PROTEIN AND KEEPS  
 CC 2 OF THE 4 MN-ATOMS ASSOCIATED WITH PSII.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: LUMENAL SURFACE OF THYLAKOID MEMBRANE.  
 CC -!- SIMILARITY: Belongs to the psbO family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; J03002; AAA87283.1; -;  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF01716; MSP; 1.  
 KW Photosynthesis; Photosystem II; Manganese; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 277 PHOTOSYSTEM II MANGANESE-STABILIZING  
 FT POLYPEPTIDE  
 FT  
 SQ SEQUENCE 277 AA; 29304 MW; 8FFC990F6D6556E CRC64;  
 Query Match 76.5%; Score 26; DB 1; Length 277;  
 Best Local Similarity 62.5%; Pred. No. 33;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVXTXIRP 9  
 DB 242 GVFTAIQ 249  
 RESULT 8  
 PANC MYCLE  
 ID -PANC MYCLE STANDARD; PRT; 313 AA.  
 AC O69524;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)  
 DE (Pantoate activating enzyme).  
 GN PANC OR MLC2548.01C.  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus."  
 RL Nature 409:1007-1011(2001).  
 CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +  
 CC diphosphate + (R)-pantothenate.  
 CC -!- PATHWAY: Pantothenate biosynthesis; last step.  
 CC -!- SIMILARITY: Belongs to the pantothenate synthetase family.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL583917; CAC29738.1; -;  
 DR EMBL; AL023093; CAA18788.1; -;  
 DR PIR; F86937; F86937.  
 DR Leproma; MF 00158; -; 1.  
 DR HAMAP; NF 00158; -; 1.  
 DR InterPro; IPR003721; Pantoate ligase.  
 DR Pfam; PF02569; Pantoate ligase; 1.  
 DR TIGRFAMs; TIGR00018; pAnc; 1.  
 KW Pantothenate biosynthesis; Ligase; Complete proteome.  
 SQ SEQUENCE 313 AA; 33585 MW; 0209C216B926A712 CRC64;  
 Query Match 76.5%; Score 26; DB 1; Length 313;  
 Best Local Similarity 50.0%; Pred. No. 37;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 GVXTXIRP 9  
 DB 116 GLRTTVP 123  
 RESULT 9  
 YD60 MYCTU  
 ID -YD60 MYCTU STANDARD; PRT; 340 AA.  
 AC Q11030;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein RV1360/MT1405/MB1395 precursor.  
 GN RV1360 OR MT1405 OR MTCY02B10.24 OR MB1395.  
 OS Mycobacterium tuberculosis, and  
 OS Mycobacterium bovis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773, 1765;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=H37Rv;  
 RX MEDLINE=9829598; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Teklaia F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;  
 RX MEDLINE=22206494; PubMed=12218036;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bissal W., Jacobs W.R. Jr, Venter J.C., Fraser C.M.;  
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL J. Bacteriol. 184:5479-5490(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=M.bovis; STRAIN=AF2122/97;  
 RX MEDLINE=22709107; PubMed=12768972;



```

ADRI_HUMAN
ID ADRI_HUMAN STANDARD; PRT; 375 AA.
AC Q96A54; Q9Y360;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Adiponectin receptor protein 1 (CGI-45).
GN ADIPOR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093;
RA Lai C.-H., Chou C.-Y., Chang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
RL Caenorhabditis elegans by comparative proteomics.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RN [3]
RP FUNCTION, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=22687101; PubMed=12802337;
RA Sugiyama T., Miyagishi M., Hara K., Tsunoda M., Murakami K.,
RA Ohteki T., Uchida S., Takekawa S., Waki H., Tsuno N.H., Shibata Y.,
RA Terauchi Y., Froquel P., Tobe K., Koyasu S., Taira K., Kitamura T.,
RA Shimizu T., Nagai R., Kadowaki T.;
RT "Cloning of adiponectin receptors that mediate antidiabetic metabolic
RL effects.";
RN [4]
RP FUNCTION: Receptor for globular and full-length adiponectin
(CC (APM1), an essential hormone secreted by adipocytes that acts as
CC an antidiabetic. Probably involved in metabolic pathways that
CC regulate lipid metabolism such as fatty acid oxidation. Mediates
CC increased AMPK, PPARα ligand activity, fatty acid oxidation and
CC glucose uptake by adiponectin. Has some high-affinity receptor for
CC globular adiponectin but low-affinity receptor for full-length
CC adiponectin.
CC
CC -!- SUBUNIT: May form homo and heteromultimers.
CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localized to the
CC cell membrane and intracellular organelles.
CC
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in skeletal
CC muscle. Expressed at intermediate level in brain, heart, spleen,
CC kidney, liver, placenta, lung and peripheral blood leukocytes.
CC Weakly expressed in colon, thymus and small intestine.
CC
CC -!- SIMILARITY: Belongs to the ADIPOR family.
CC
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 369.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF151803; AAD34040.1; ALT_FRAME.
CC EMBL; BC001594; AAH01594.1; -.
CC EMBL; BC010743; AAH10743.1; -.
CC MIM; 607945; -.
CC InterPro; IPR004254; HlyIII-related.
CC Pfam; PF03006; UPF0073; 1.
CC KW Fatty acid metabolism; Lipid metabolism; Receptor; Transmembrane.
CC FT DOMAIN 1 136 157
CC FT TRANSMEM 137 157
CC FT DOMAIN 158 170
CC FT TRANSMEM 171 191
CC FT DOMAIN 192 209
CC FT TRANSMEM 210 230
CC FT DOMAIN 231 235
CC FT TRANSMEM 236 256
CC FT DOMAIN 257 267
CC FT TRANSMEM 268 288
CC FT DOMAIN 289 296
CC FT TRANSMEM 297 317
CC FT DOMAIN 318 331
CC FT TRANSMEM 332 352
CC FT DOMAIN 353 375
CC FT SEQUENCE 375 AA; 42615 MW; 1CC0300A7D178EB0 CRC64;
CC
CC Query Match 76.5%; Score 26; DB 1; Length 375;
CC Best Local Similarity 50.0%; Pred. No. 44;
CC Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 2 GVXTXIRP 9
CC Db 152 GILTMWRP 159
CC
CC RESULT 12
CC ADRI_MOUSE STANDARD; PRT; 375 AA.
CC ID ADRI_MOUSE STANDARD; PRT; 375 AA.
CC AC Q91VH1; Q9CZAO;
CC DT 15-MAR-2004 (Rel. 43, Created)
CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Adiponectin receptor protein 1.
CC GN ADIPOR1.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=C57BL/6J; TISSUE=Embryo;
CC RX MEDLINE=21085660; PubMed=11217851;
CC RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
CC RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
CC RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
CC RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
CC RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
CC RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
CC RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
CC RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
CC RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
CC RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
CC RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
CC RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
CC RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
CC RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
CC RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
CC RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

```

PT	TRANSMEM	137	157	1 (POTENTIAL). EXTRACELLULAR (POTENTIAL).
PT	DOMAIN	158	170	2 (POTENTIAL).
PT	TRANSMEM	171	191	3 (POTENTIAL).
PT	DOMAIN	192	209	4 (POTENTIAL).
PT	TRANSMEM	210	230	5 (POTENTIAL).
PT	DOMAIN	231	235	6 (POTENTIAL).
PT	TRANSMEM	236	256	7 (POTENTIAL).
PT	DOMAIN	257	267	8 (POTENTIAL).
PT	TRANSMEM	268	288	9 (POTENTIAL).
PT	DOMAIN	289	296	10 (POTENTIAL).
PT	TRANSMEM	297	317	11 (POTENTIAL).
PT	DOMAIN	318	331	12 (POTENTIAL).
PT	TRANSMEM	332	352	13 (POTENTIAL).
PT	DOMAIN	353	375	14 (POTENTIAL).
SEQ	SEQUENCE	375 AA; 42366 MW; OE72F81B5E938CE CRC64;		

Query Match 76.5%; Score 26; DB 1; Length 375;  
 Best Local Similarity 50.0%; Pred. No. 44;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy	2 GVXIXIRP 9
Db	152 GILTMWRP 159

1: |||

RESULT 13

ID	YMN3_PODAN	STANDARD;	PRT; 433 AA.
AC	P15563;		
DT	01-APR-1990 (Rel. 14, Created)		
DT	01-APR-1990 (Rel. 14, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Hypothetical 49.1 kDa protein in ND3 intron.		
OS	Podospora anserina.		
CG	Mitochondrion.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
OC	Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.		
CN	NCBI_TaxID=5145;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=s, and A;		
RX	MEDLINE=89125610; PubMed=2975708;		
RA	Cummings D.J., Domenico J.M.;		
RT	"Sequence analysis of mitochondrial DNA from Podospora anserina.		
RT	Pervasiveness of a class I intron in three separate genes.";		
RL	J. Mol. Biol. 204:815-839(1988).		
RP	[2]		
RN	COMPLETE GENOME.		
RC	STRAIN=s;		
RX	MEDLINE=90291512; PubMed=2357736;		
RA	Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;		
RT	"The complete DNA sequence of the mitochondrial genome of Podospora		
RT	anserina.";		
RL	Curr. Genet. 17:375-402(1990).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; X55026; CAA38767.2; -		
DR	EMBL; X14485; -; NOT ANNOTATED_CDS.		
DR	PIR; S05654; S05654.		
DR	InterPro; IPR001982; Endonuc LAG/HNH.		
DR	Pfam; PF00961; LAGLIDADG_1; 2.		
KW	Hypothetical protein; Mitochondrion.		
SEQ	SEQUENCE 433 AA; 49089 MW; 75139322A8B477D0 CRC64;		

Query Match 76.5%; Score 26; DB 1; Length 433;  
 Best Local Similarity 50.0%; Pred. No. 50;

Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVATXIRP 9  
Db 284 GINTILRP 291

RESULT 14  
VL2\_HPV19  
ID\_VL2\_HPV19 STANDARD; PRT; 520 AA.  
AC P36752;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Minor capsid protein L2.  
GN L2.  
OS Human papillomavirus type 19.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10608;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9426501; PubMed=8205638;  
RA Delius H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X74470; CAA52522.1; --  
DR PUB; S36489; S36489  
DR InterPro; IPR000784; Late\_L2.  
DR Pfam; PF00513; Late\_protein\_L2; 1.  
KW Coat protein; Late protein.  
SQ SEQUENCE 520 AA; 57019 MW; F428B9097A981AE1 CRC64;  
  
Query Match 76.5%; Score 26; DB 1; Length 520;  
Best Local Similarity 62.5%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 GVATXIRP 9  
Db 81 GTATVIRP 88

RESULT 15  
RYR2\_HUMAN  
ID\_RYR2\_HUMAN STANDARD; PRT; 4967 AA.  
AC Q92736; Q15411;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RYR2)  
DE (RYR-2) (Cardiac muscle ryanodine receptor-calcium release channel)  
GN RYR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
RC TISSUE=Heart muscle;  
RX MEDLINE=96404895; PubMed=8809036;  
RA Tunwell R.E.A., Wickenden C., Bertrand B.M.A., Shevchenko V.I.,  
RA Walsh M.B., Allen P.D., Lai F.A.;  
RT "The human cardiac muscle ryanodine receptor-calcium release channel:

RT identification, primary structure and topological analysis.";  
RL Biochem. J. 318:477-487(1996).  
RN [2]  
RP SEQUENCE OF 9-87 AND 533-681 FROM N.A., DEVELOPMENTAL STAGE, AND  
RP INDUCTION.  
RC TISSUE=Heart muscle, and Myometrium;  
RX MEDLINE=97220386; PubMed=9148749;  
RA Awad S.S., Lamb H.K., Morgan J.M., Dunlop W., Gillespie J.I.;  
RT "Differential expression of ryanodine receptor RYR2 mRNA in the non-  
RT pregnant and pregnant human myometrium.";  
RL Biochem. J. 322:777-783(1997).  
RN [3]  
RP SEQUENCE OF 4292-4479 FROM N.A., AND TISSUE SPECIFICITY.  
RC TISSUE=Cerebellum, and Hippocampus;  
RX MEDLINE=98268728; PubMed=9607712;  
RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;  
RT "Partial cloning and differential expression of ryanodine  
RT receptor/calcium-release channel genes in human tissues including the  
RT hippocampus and cerebellum.";  
RL Neuroscience 85:205-216(1998).  
RN [4]  
RP VARIANTS VTSIP LEU-2246; SER-2474; LYS-4104 AND CVS-4497.  
RX PubMed=11208676;  
RA Priori S.G., Napolitano C., Tiso N., Memmi M., Vignati G., Bloise R.,  
RA Sorrentino V.V., Danieli G.A.;  
RT "Mutations in the cardiac ryanodine receptor gene (hRYR2) underlie  
RT catecholaminergic polymorphic ventricular tachycardia.";  
RL Circulation 103:196-200(2001).  
RN [5]  
RP VARIANTS FPVT SER-2328; ARG-4201 AND PHE-4653, AND VARIANT ARG-2958.  
RX MEDLINE=21112862; PubMed=11157710;  
RA Laitinen P.J., Brown K.M., Piippo K., Swan H., Devaney J.M.,  
RA Brahmabhatt B., Donarum E.A., Marino M., Tiso N., Vaitasalo M.,  
RA Toivonen L., Stephan D.A., Kontula K.;  
RT "Mutations of the cardiac ryanodine receptor (RYR2) gene in familial  
RT polymorphic ventricular tachycardia.";  
RL Circulation 103:485-490(2001).  
RN [6]  
RP VARIANTS ARVD2 PRO-433 AND ILE-2386.  
RX MEDLINE=21096894; PubMed=11159936;  
RA Tiso N., Stephan D.A., Nava A., Bagattin A., Devaney J.M., Stanchi F.,  
RA Larderet G., Brahmabhatt B., Brown K., Baucé B., Mariago M., Basso C.,  
RA Thiene G., Danieli G.A., Rampazzo A.;  
RT "Identification of mutations in the cardiac ryanodine receptor gene in  
RT families affected with arrhythmic right ventricular  
RT cardiomyopathy type 2 (ARVD2).";  
RL Hum. Mol. Genet. 10:189-194(2001).  
CC -!- FUNCTION: Communication between transverse-tubules and  
CC sarcoplasmic reticulum. Contraction of cardiac muscle is triggered  
CC by release of calcium ions from SR following depolarization of T-  
CC tubules (By similarity).  
CC -!- SUBUNIT: Homotetramer (Potential).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing, Named isoforms=2;  
CC Name=1;  
CC IsoId=Q92736-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q92736-2; Sequence=VSP\_005953;  
CC -!- TISSUE SPECIFICITY: Heart muscle, brain (cerebellum and  
CC hippocampus) and placenta.  
CC -!- DEVELOPMENTAL STAGE: Expressed in myometrium during pregnancy.  
CC -!- INDUCTION: By TGF-beta.  
CC -!- DISEASE: Defects in RYR2 are the cause of arrhythmic right  
CC ventricular cardiomyopathy type 2 (ARVD2) [MIM:600996]. ARVD2 is  
CC an autosomal dominant disease characterized by partial  
CC degeneration of the myocardium of the right ventricle, electrical  
CC instability, and sudden death. It is clinically defined by  
CC electrocardiographic and angiographic criteria; pathologic  
CC findings, replacement of ventricular myocardium with fatty and  
CC fibrous elements, preferentially involve the right ventricular  
CC free wall.  
CC -!- DISEASE: Defects in RYR2 are the cause of an autosomal dominant

form of stress-induced polymorphic ventricular tachycardia (VTSP) [MIM:604772]; also known as catecholaminergic polymorphic ventricular tachycardia. VTSP is a genetic arrhythmogenic disorder characterized by stress-induced, bidirectional ventricular tachycardia that may degenerate into cardiac arrest and cause sudden death.

-1- DISEASE: Defects in RYR2 are a cause of familial polymorphic ventricular tachycardia (FPVT) [MIM:192605]. FPVT is an autosomal-dominant, inherited disease with a relatively early onset and a mortality rate of approximately 30% by the age of 30 years. Phenotypically, it is characterized by salvos of bidirectional and polymorphic ventricular tachycardias in response to vigorous exercise, with no structural evidence of myocardial disease.

-1- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-1- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-1- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-1- SIMILARITY: Belongs to the ryanodine receptor family.

-1- SIMILARITY: Contains 5 MIR domains.

-1- SIMILARITY: Contains 3 SPRY domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X98330; CAA66975.1; -  
 DR EMBL; Y08218; CAA69395.1; -  
 DR EMBL; X91869; CAA62975.1; -  
 DR EMBL; AJ002511; CAA05502.1; -  
 DR FIR; S72289; S72269  
 DR Genew; HGNC:10484; RYR2.  
 DR MIM; 180902; -  
 DR MIM; 600996; -  
 DR MIM; 604772; -  
 DR MIM; 192605; -  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0015278; P:calcium-release channel activity; TAS.  
 DR GO; GO:0006938; P:muscle contraction; TAS.  
 DR GO; GO:0008016; P:regulation of heart rate; TAS.  
 DR InterPro; IPR000699; Ca-rel\_channel.  
 DR InterPro; IPR001682; Ca/Na\_pore.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR003608; MIR.  
 DR InterPro; IPR001215; Ryanodn\_receptor.  
 DR InterPro; IPR003032; RYR.  
 DR InterPro; IPR003877; SPRY\_receptor.  
 DR Pfam; PF00036; ehand; 2.  
 DR Pfam; PF0520; ion\_trans; 1.  
 DR Pfam; PF02815; MIR; 4.  
 DR Pfam; PF01365; RYDR\_ITPR; 2.  
 DR Pfam; PF02026; RYR; 4.  
 DR Pfam; PF00622; SPRY; 3.  
 DR PRINTS; PR00795; RYANODINER.  
 DR SMART; SM00054; EFH; 2.  
 DR SMART; SM00472; MIR; 4.  
 DR SMART; SM00449; SPRY; 3.  
 DR PROSITE; PS50919; MIR; 5.  
 KW Ionic channel; Receptor; Calcium channel; Calmodulin-binding; Repeat;  
 KW Transmembrane; Glycoprotein; Phosphorylation; Alternative splicing;  
 KW Polymorphism; Disease mutation; Cardiomyopathy; Multigene family.  
 FT DOMAIN 1 3089 CYTOPLASMIC (BY SIMILARITY).  
 FT TRANSMEM 3090 3109 M' (POTENTIAL).

FT TRANSMEM 3153 3171 M'' (POTENTIAL).  
 FT TRANSMEM 3940 3959 M1 (POTENTIAL).  
 FT TRANSMEM 3978 3995 M2 (POTENTIAL).  
 FT TRANSMEM 4233 4256 M3 (POTENTIAL).  
 FT TRANSMEM 4294 4314 M4 (POTENTIAL).  
 FT TRANSMEM 4500 4520 M5 (POTENTIAL).  
 FT TRANSMEM 4578 4600 M6 (POTENTIAL).  
 FT TRANSMEM 4720 4740 M7 (POTENTIAL).  
 FT TRANSMEM 4768 4786 M8 (POTENTIAL).  
 FT TRANSMEM 4810 4827 M9 (POTENTIAL).  
 FT TRANSMEM 4845 4867 M10 (POTENTIAL).  
 FT DOMAIN 110 165 MIR 1.  
 FT DOMAIN 172 217 MIR 2.  
 FT DOMAIN 225 280 MIR 3.  
 FT DOMAIN 286 343 MIR 4.  
 FT DOMAIN 351 408 MIR 5.  
 FT DOMAIN 670 808 SPRY 1.  
 FT DOMAIN 1098 1221 SPRY 2.  
 FT DOMAIN 1423 1561 SPRY 3.  
 FT DOMAIN 4414 4455 GLU-RICH (ACIDIC).  
 FT DOMAIN 853 2925 4 X APPROXIMATE REPEATS.  
 FT REPEAT 967 1080 1.  
 FT REPEAT 292 2810 2.  
 FT REPEAT 2812 2825 3.  
 FT BINDING 2618 3015 MODULATOR (POTENTIAL).  
 FT BINDING 2774 2806 CALMODULIN (POTENTIAL).  
 FT BINDING 2876 2897 CALMODULIN (POTENTIAL).  
 FT BINDING 2997 3015 PHOSPHORYLATION (BY CAMK) (BY SIMILARITY).  
 FT MOD\_RES 2808 2808 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3427 3427 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3536 3536 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3770 3770 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 76.5%; Score 26; DB 1; Length 4967;

Best Local Similarity 50.0%; Pred. No. 5.3e+02; Mismatches 2; Indels 0; Gaps 0;

Matches 4; Conservative 2;

QY 2 GVXTXIRP 9

Db 1752 GLSTSLRP 1759

Search completed: April 1, 2004, 17:39:26

Job time : 5.52632 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 28.4211 seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-5

Perfect score: 34

Sequence: 1 XGVXTXIRP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext. 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriopl:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	91.2	366	12 Q91TT6	Q91tt6 tupaiia herp
2	31	91.2	377	16 Q9PFS8	Q9pfs8 ureaplasma
3	31	91.2	516	10 Q94D96	Q94d96 oryza sativ
4	30	88.2	96	10 Q8S0Y8	Q8s0y8 oryza sativ
5	30	88.2	139	2 Q84CK2	Q84ck2 streptomyc
6	30	88.2	156	10 Q7X6K7	Q7x6k7 oryza sativ
7	30	88.2	2796	2 Q48926	Q48926 mycobacteri
8	29	85.3	185	16 Q9F2Y6	Q9f2y6 streptomyc
9	29	85.3	312	16 Q82HK0	Q82hk0 streptomyc
10	29	85.3	314	16 Q8NLF3	Q8nlf3 corynebacte
11	29	85.3	407	16 Q8F364	Q8f364 leptospira
12	29	85.3	464	16 Q89QRO	Q89qro bradyrhizob
13	29	85.3	502	16 Q7WPH7	Q7wph7 bordetella
14	29	85.3	619	2 Q83LL7	Q83ll7 salmonella
15	29	85.3	5035	6 Q29104	Q29104 sus scrofa
16	29	85.3	5035	6 Q29105	Q29105 sus scrofa

Q80x16 mus musculu  
Q9h92 halobacteri  
Q07338 bacillus su  
P96566 amycolatops  
Q91108 oryza sativ  
Q8r946 thelmoanaer  
Q22659 caenorhabdi  
Q85716 mycobacteri  
Q98qf6 mycoplasma  
Q82pr2 streptomyc  
Q9kdg9 bacillus ha  
Q9cr40 mus musculu  
Q87h35 vibrio para  
Q8iue3 homo sapien  
Q9l214 streptomyc  
Q2zx98 bacterioph  
Q39919 human cytom  
O66312 unidentified  
Q8zwe4 pyrobaculum  
Q8u0f4 pyrococcus  
Q8pgv5 xanthomonas  
Q88y76 lactobacill  
Q9f620 rhizobium m  
Q92rx1 rhizobium m  
Q97da0 clostridium  
Q7zy04 xenopus lae  
Q9a597 caulobacter  
Q81u14 bacillus an  
Q8vpp4 micrococcus

## ALIGNMENTS

### RESULT 1

Q91TT6 PRELIMINARY; PRT; 366 AA.  
ID Q91TT6  
AC Q91TT6; 01-DEC-2001 (TREMELrel. 19, Created)  
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)  
DE T22.3.  
OS Tupaia herpesvirus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae.  
OX NCBI\_TaxID=10397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RX MEDLINE=21211637; PubMed=11312357;  
RA Bahr U., Darai G.;  
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree Shrew) Herpesvirus."  
RL J. Virol. 75:4854-4870(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RA Darai G., Bahr U.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF281817; AAK57051.1; -.  
SQ SEQUENCE 366 AA; 42795 MW; 757ED66A38D827FE CRC64;  
Query Match 91.2%; Score 31; DB 12; Length 366;  
Best Local Similarity 75.0%; Pred. NO. 21;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9

Db 162 GVSTSRP 169

### RESULT 2

Q9PFS8



```

ID Q9PPS8 PRELIMINARY; PRT; 377 AA.
AC Q9PPS8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Oligopeptide transport system permease protein.
GN OPPB OR U0562.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RL Nature 407:757-762(2000).
DR GO; AB002154; AAF30975.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp; 1.
KW Complete proteome.
SQ SEQUENCE 377 AA; 41923 MW; 391626240B625A6F CRC64;

Query Match 91.2%; Score 31; DB 16; Length 377;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 146 GVSTXIRP 153

RESULT 3
Q94D96 PRELIMINARY; PRT; 516 AA.
ID Q94D96;
AC Q94D96;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P0439E11.9 protein.
GN P0439E11.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0439E11."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003315; BAB61259.1; --
DR EMBL; AP003315; BAB61259.1; --
DR Gramene; Q94D96; --
DR InterPro; IPR005529; P:sugar binding; IEA.
DR InterPro; IPR001480; B:lectin.
DR Pfam; PF01453; Agglutinin; 1.
DR SMART; SM00108; B:lectin; 1.
SQ SEQUENCE 516 AA; 55483 MW; CB9B88B9E66377700 CRC64;

Query Match 91.2%; Score 31; DB 10; Length 516;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 349 GVTAIRP 356

```

```

RESULT 4
Q8SOY8 PRELIMINARY; PRT; 96 AA.
ID Q8SOY8;
AC Q8SOY8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 22, Last annotation update)
DE P0485B12.23 protein (P0031D02.2 protein).
GN P0485B12.29 OR P0031D02.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
RT clone: P0485B12."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003348; BAB86498.1; --
DR EMBL; AP003348; BAB86498.1; --
DR Gramene; Q8SOY8; --
DR InterPro; IPR009002; FWN binding.
SQ SEQUENCE 96 AA; 10908 MW; 2F93BF8513818BF8 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 96;
Best Local Similarity 62.5%; Pred. No. 9.7;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 67 GVTSVRP 74

RESULT 5
Q84CK2 PRELIMINARY; PRT; 139 AA.
ID Q84CK2;
AC Q84CK2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kinul.
GN KINUL.
OS Streptomyces murayamaensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=224537;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin X.H., Mahadevan B., Grochowski L., Proteau P.J.;
RT "Molecular cloning and sequence of the kinamycin angucycline type II
RT polyketide synthase gene cluster from Streptomyces murayamaensis."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY228175; AAO65341.1; --
DR InterPro; IPR009002; FWN binding.
SQ SEQUENCE 139 AA; 14815 MW; 65A27123A80349E4 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 139;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 111

```

```
Db 28 GVLTIRP 35

RESULT 6
Q7X6K7 PRELIMINARY; PRT; 156 AA.
ID Q7X6K7
AC Q7X6K7 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE OSUNBA0018J19.23 protein [OSUNBA0034E24.1 protein].
DE OSUNBA0018J19.23 OR OSUNBA0034E24.1.
GN Oryza sativa (Rice).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OC NCBI_TaxID=4530;
RN [1]
RN SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hae P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AL662956; CAE04456.1; -
DR EMBL; AL662957; CAE05907.1; -
SQ SEQUENCE 156 AA; 17088 MW; C8B47B96DFB7BD02 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 156;
Best Local Similarity 62.5%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 30 GVTSVRP 37

RESULT 7
Q48926 PRELIMINARY; PRT; 2796 AA.
ID Q48926
AC Q48926
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fatty acid synthase.
DE FAS.
GN FAS.
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OC NCBI_TaxID=1765;
RN [1]
RN SEQUENCE FROM N.A.
RA STRAIN=BCG-Pasteur;
RA MEDLINE=96200863; PubMed=8621098;
RA Fernandes N.D., Kolattukudy P.E.;
RA "Cloning, sequencing and characterization of a fatty acid synthase-
RA encoding gene from Mycobacterium tuberculosis var. bovis BCG.";
RL Gene 170:95-99 (1996).
DR EMBL; U36763; AA030809.1; -
DR PIR; JC4743; JC4743.
DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac trans.
DR InterPro; IPR003965; Fatty_acid_synth.

DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR002539; MacC_dehydratas.
DR Pfam; PF00698; Acyl_transf; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; ketoacyl-synt_C; 1.
DR Pfam; PF01575; MacC_dehydratas; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
SQ SEQUENCE 2796 AA; 303675 MW; 47B87169126D2504 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 2796;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 2704 GVTSIRP 2711

RESULT 8
Q9F2Y6 PRELIMINARY; PRT; 185 AA.
ID Q9F2Y6
AC Q9F2Y6 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative integral membrane protein.
DE SCO4361 OR SCD19.16.
GN Streptomyces coelicolor.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
RN [1]
RN SEQUENCE FROM N.A.
RA STRAIN=A3(2) / M145;
RA MEDLINE=21396410; PubMed=12000553;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kisser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kisser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RA coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL939119; CAC08317.1; -
KN Complete proteome.
SQ SEQUENCE 185 AA; 19332 MW; E2FE76A5D21296FD CRC64;

Query Match 85.3%; Score 29; DB 16; Length 185;
Best Local Similarity 75.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 73 GVTSIRP 80

RESULT 9
Q82HKO PRELIMINARY; PRT; 312 AA.
ID Q82HKO
AC Q82HKO (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative secreted substrate-binding protein.
DE SAV3508.
GN Streptomyces avermitilis.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
```

```

NCBI_TaxID=33903;
[1]
SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
[2]
SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005035; BAC71220.1; -.
DR InterPro; IPR001092; HLH basic.
DR InterPro; IPR007210; OpuA.
DR Pfam; PF04069; OpuA; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Complete proteome.
SQ SEQUENCE 312 AA; 32672 MW; CADA8070DA5C2210 CRC64;

Query Match 85.3%; Score 29; DB 16; Length 312;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 71 GVRTVRP 78

RESULT 10
Q8NLF3 PRELIMINARY; PRT; 314 AA.
ID Q8NLF3
AC Q8NLF3
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein Cgl2989.
GN CGL2989.
OS Corynebacterium glutamicum (Bravibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
[1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP005283; BAC00383.1; -.
DR GO; GO:0006950; P:response to stress; IEA.
DR InterPro; IPR006016; Usp_dom.
DR Pfam; PF00582; Usp; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 314 AA; 34512 MW; B2423836E2E4DBDC CRC64;

Query Match 85.3%; Score 29; DB 16; Length 314;
Best Local Similarity 62.5%; Pred. No. 57;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 46 GIVTIRP 53

```

```

RESULT 11
Q8F364 PRELIMINARY; PRT; 407 AA.
ID Q8F364
AC Q8F364;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capb protein, Mur ligase family.
GN CAPB OR LA2547.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiaceae; Leptospira.
OX NCBI_TaxID=173;
[1]
SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB011422; AAN49746.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0015874; F:ligase activity; IEA.
DR GO; GO:0045227; P:capsule polysaccharide biosynthesis; IEA.
DR InterPro; IPR008337; CapB.
DR InterPro; IPR001005; MYD_DNA_binding.
DR PRINTS; PR01758; CAPSULEPROTB.
DR PROSITE; PS00037; MYB_1; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 407 AA; 46759 MW; AF12FAB606746B4B CRC64;

Query Match 85.3%; Score 29; DB 16; Length 407;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9
DB 132 GVITNRP 139

RESULT 12
Q89QR0 PRELIMINARY; PRT; 464 AA.
ID Q89QR0
AC Q89QR0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Succinyl-diaminopimelate desuccinylase.
GN DAPE OR BLR3064.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasanoto S., Matsumoto A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005946; BAC48329.1; -.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
KW Complete proteome.
SQ SEQUENCE 464 AA; 50583 MW; B12C32D8F9432CA1 CRC64;

```

Query Match 85.3%; Score 29; DB 16; Length 464;  
Best Local Similarity 52.5%; Pred. No. 95;  
Matches 5; Conservative 1; Mismatches 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||  
Db 58 GVATEVRP 65

RESULT 13  
Q7WPH7 PRELIMINARY; PRT; 502 AA.  
AC Q7WPH7  
DT 01-OCT-2003 (TREMELrel. 25, Created)  
DT 01-OCT-2003 (TREMELrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Probable GntR-family transcriptional regulator.  
GN BB0707.  
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Alcaligenaceae; Bordetella.  
OX NCBI\_TaxID=518;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RB50 / ATCC BAA-588;  
RX MEDLINE=22827954; PubMed=12910271;  
RA Parthill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
RA Achten-Tarrega A.M., Temple L., James K., Harris B., Quail M.A.,  
RA Achman M., Atkin R., Baker S., Basham D., Barton A., Davis P., Cherevach I.,  
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
RA Leathem S., Moulé S., Norberczak H., O'Neil S., Ormond D., Price C.,  
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
RA Sharp S., Simmonds M., Skelton J., Squares R., Stevens K.,  
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,  
RA "Comparative analysis of the genome sequences of Bordetella pertussis,  
RT Bordetella parapertussis and Bordetella bronchiseptica.",  
RL Nat. Genet. 35:32-40(2003).  
DR EMBL; BX640439; CAE31207.1; -.  
KW Complete proteome.  
SQ SEQUENCE 502 AA; 54866 MW; 19C15B8F2F34EB8D CRC64;

Query Match 85.3%; Score 29; DB 16; Length 502;  
Best Local Similarity 50.0%; Pred. No. 92;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||  
Db 424 GIATAVRP 431

RESULT 14  
Q8RL7 PRELIMINARY; PRT; 619 AA.  
AC Q8RL7  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Salmonella enteritidis.  
OG Plasmid pp.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=592;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Rychlik I., Gregorova D., Doskocilova J., Sebkova A.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY079199; AL87010.1; -.  
DR GO; GO:0046621; C:extrachromosomal DNA; IEA.  
DR GO; GO:0009291; P:unidirectional conjugation; IEA.  
DR InterPro; IPR005053; MobA\_MobL.

DR Pfam; PF03389; MobA\_MobL; 1.  
KW Hypothetical protein; Plasmid.  
SQ SEQUENCE 619 AA; 71857 MW; BF15PFS01426DAF5 CRC64;

Query Match 85.3%; Score 29; DB 2; Length 619;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||  
Db 528 GVKTEIRP 535

RESULT 15  
Q29104 PRELIMINARY; PRT; 5035 AA.  
AC Q29104  
DT 01-NOV-1996 (TREMELrel. 01, Created)  
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Ryanodine receptor.  
GN RYR1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SPRAIN=Belgisch Landras/Pietrain; TISSUE=Skeletal muscle;  
RX MEDLINE=91320118; PubMed=1862346;  
RA Fujii J., Otsu K., Zorzato F., de Leon S., Khanna V.K., Weiler J.E.,  
RA O'Brien P.J., MacLennan D.H.;  
RT "Identification of a mutation in porcine ryanodine receptor associated  
with malignant hyperthermia.",  
RL Science 253:448-451(1991).  
DR EMBL; M91451; AAA31118.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005262; F:calcium channel activity; IEA.  
DR GO; GO:0005509; F:calcium ion binding; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0006812; P:calcium ion transport; IEA.  
DR GO; GO:0006812; P:cation transport; IEA.  
DR InterPro; IPR000699; Ca-rel\_channel.  
DR InterPro; IPR001682; Ca/Na\_Pore.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR005821; Ion\_trans.  
DR InterPro; IPR003608; MIR.  
DR InterPro; IPR003877; SPRY\_receptor.  
DR InterPro; IPR003032; RYR.  
DR InterPro; IPR001215; Ryanodn\_receptor.  
DR Pfam; PF00520; ion\_trans\_1.  
DR Pfam; PF02815; MIR\_4.  
DR Pfam; PF01385; RYR\_ITPR\_2.  
DR Pfam; PF02026; RYR\_4.  
DR Pfam; PF00622; SPRY\_3.  
DR PRINTS; PR00795; RYANODINER.  
DR SMART; SM00472; MIR\_4.  
DR SMART; SM00449; SPRY\_3.  
KW Ionic channel; Receptor; Transmembrane.  
SQ SEQUENCE 5035 AA; 565089 MW; 66B0E1AA590636EF CRC64;

Query Match 85.3%; Score 29; DB 6; Length 5035;  
Best Local Similarity 62.5%; Pred. No. 9.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 GVXTXIRP 9  
|||  
Db 1767 GVITSLRP 1774

Search completed: April 1, 2004, 17:38:42  
Job time : 29.4211 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:25:27 ; Search time 52.6842 Seconds  
(without alignments)  
58.993 Million cell updates/sec

Title: US-09-833-196-6

Perfect score: 32

Sequence: 1 XGVVXXTRPX 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	78.1	8	7 ADD31283	Add31283 Angiogene
2	25	78.1	8	7 ADD31284	Add31284 Angiogene
3	25	78.1	8	7 ADD31285	Add31285 Angiogene
4	25	78.1	8	7 ADE83634	Ade83634 Antiangi
5	25	78.1	8	7 ADE83636	Ade83636 Antiangi
6	25	78.1	8	7 ADE83635	Ade83635 Antiangi
7	25	78.1	9	3 RAY67983	Ray67983 Antiangi
8	25	78.1	9	4 AAU15042	Aau15042 Antiangi
9	25	78.1	9	4 AAU15043	Aau15043 Antiangi
10	25	78.1	9	4 AAU15044	Aau15044 Antiangi
11	25	78.1	9	4 AAU15027	Aau15027 Antiangi
12	25	78.1	9	4 AAU15041	Aau15041 Antiangi
13	25	78.1	9	4 AAU15016	Aau15016 Antiangi
14	25	78.1	9	6 AAC26627	Aac26627 Anti-angi
15	25	78.1	9	6 AAC26633	Aac26633 Anti-angi
16	25	78.1	9	6 AAC26637	Aac26637 Anti-angi
17	25	78.1	9	7 ADD69919	Add69919 Antiangi
18	25	78.1	10	4 AAU15050	Aau15050 Antiangi
19	25	78.1	10	4 AAU14987	Aau14987 Antiangi
20	25	78.1	10	4 AAU14993	Aau14993 Antiangi
21	25	78.1	10	6 AAO26644	Aao26644 Anti-angi
22	25	78.1	136	3 RAG35938	Rag35938 Zea mays
23	25	78.1	139	6 APE57616	Ape57616 S. muraya
24	25	78.1	505	3 AAG41711	Aag41711 Arabidops
25	25	78.1	540	3 AAG41710	Aag41710 Arabidops

26	25	78.1	561	3 AAG41709	Aag41709 Arabidops
27	25	78.1	1061	2 AAW87504	Aaw87504 Human N-m
28	25	78.1	1081	3 AAB36240	Aab36240 Human N-m
29	25	78.1	1081	5 ABB56523	Abb56523 Human NMD
30	25	78.1	1081	6 ABU61440	Abu61440 Human N-m
31	25	78.1	1081	7 AAE39251	Aae39251 Human NMD
32	25	78.1	1212	2 AAW87503	Aaw87503 Human N-m
33	25	78.1	1232	3 AAB26239	Aab26239 Human NMD
34	25	78.1	1232	5 AAM47961	Aam47961 Human NMD
35	25	78.1	1232	6 ABU61439	Abu61439 Human NMD
36	25	78.1	1232	7 AAE39250	Aae39250 Human NMD
37	25	78.1	2796	6 ABU79133	Abu79133 Mycobacte
38	24	75.0	8	7 ADD31274	Add31274 Angiogene
39	24	75.0	8	7 ADD31291	Add31291 Angiogene
40	24	75.0	8	7 ADD31265	Add31265 Angiogene
41	24	75.0	8	7 ADD31289	Add31289 Angiogene
42	24	75.0	8	7 ADD31264	Add31264 Angiogene
43	24	75.0	8	7 ADD31266	Add31266 Angiogene
44	24	75.0	8	7 ADD31294	Add31294 Angiogene
45	24	75.0	8	7 ADD31286	Add31286 Angiogene

## ALIGNMENTS

RESULT 1  
ADD31283  
ID ADD31283 standard; peptide; 8 AA.  
XX AC ADD31283;  
XX

DT 15-JAN-2004 (first entry)

XX DE Angiogenesis inhibiting peptide #58.

KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antithrombotic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX OS Synthetic.  
XX

Key Location/Qualifiers  
FH Modified-site 1  
FT /label= OTHER  
FT /note= "OTHER= N-acetyl"  
FT Modified-site 3  
FT /label= OTHER  
FT Modified-site 8  
FT /note= "OTHER= D-form residue, aile (not defined)"  
FT /note= "NHCH2CH3"  
US2003109455-A1.  
PD 12-JUN-2003.  
XX  
XX 30-OCT-2002; 2002US-00283550.  
XX  
XX 31-OCT-2001; 2001US-0335017P.  
XX (HAYI/) HAYIV F.  
XX (BRAD/) BRADLEY M F.  
XX Haviv F, Bradley MP;  
XX WPI; 2003-843101/78.  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX Claim 13; Page 25; 26pp; English.

XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
CC has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
CC antiarthritic activities. (I) or a salt of it, is used in a  
CC pharmaceutical composition for inhibiting angiogenesis or for treating  
CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
CC and degenerative arthritis. This is the amino acid sequence of an  
CC angiogenesis inhibiting peptide of the invention.  
XX Sequence 8 AA;  
SQ

Query Match 78.1%; Score 25; DB 7; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
Db 1 GVXSSIRP 8  
||| |||  
||| |||

RESULT 2  
ADD31284  
ID ADD31284 standard; peptide; 8 AA.  
XX  
AC ADD31284;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Angiogenesis inhibiting peptide #59.  
XX  
KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX  
OS Synthetic.  
XX

Key Location/Qualifiers  
FH Modified-site 1 /label= OTHER  
FT Modified-site 3 /note= "N-acetyl"  
FT Modified-site 3 /label= OTHER  
FT Modified-site 8 /label= OTHER  
FT Modified-site 8 /note= "OTHER= D-form residue, alle (not defined)"  
FT Modified-site 8 /note= "NHCH2CH3"  
XX  
XX US2003109455-A1.  
XX  
XX 12-JUN-2003.  
XX  
XX 30-OCT-2002; 2002US-00283550.  
XX  
XX 31-OCT-2001; 2001US-0335017P.  
XX  
XX (HAVI/) HAVIV F.  
XX (BRAD/) BRADLEY M F.  
XX  
XX Haviv F, Bradley MF;  
XX WPI; 2003-843101/78.  
XX  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX  
XX Claim 13; Page 25; 26pp; English.  
XX  
XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
XX has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
XX antiarthritic activities. (I) or a salt of it, is used in a  
XX pharmaceutical composition for inhibiting angiogenesis or for treating  
XX cancer in a mammal. (I) Can be also used for the treatment or prevention  
XX of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
XX and degenerative arthritis. This is the amino acid sequence of an  
XX angiogenesis inhibiting peptide of the invention.  
XX Sequence 8 AA;  
SQ

CC pharmaceutical composition for inhibiting angiogenesis or for treating  
CC cancer in a mammal. (I) Can be also used for the treatment or prevention  
CC of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
CC and degenerative arthritis. This is the amino acid sequence of an  
CC angiogenesis inhibiting peptide of the invention.  
XX  
XX Sequence 8 AA;  
SQ

Query Match 78.1%; Score 25; DB 7; Length 8;  
Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
Db 1 GVXTSIRP 8  
||| |||  
||| |||

RESULT 3  
ADD31285  
ID ADD31285 standard; peptide; 8 AA.  
XX  
AC ADD31285;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Angiogenesis inhibiting peptide #60.  
XX  
KW heptapeptide compound; octapeptide compound; nonapeptide compound;  
KW antiangiogenic; cytostatic; immunosuppressive; antirheumatic;  
KW antiarthritic activities; angiogenesis inhibiting; cancer;  
KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
KW degenerative arthritis.  
XX  
OS Synthetic.  
XX

Key Location/Qualifiers  
FH Modified-site 1 /label= OTHER  
FT Modified-site 3 /note= "N-acetyl"  
FT Modified-site 3 /label= OTHER  
FT Modified-site 8 /note= "OTHER= D-form residue, alle (not defined)"  
FT Modified-site 8 /note= "NHCH2CH3"  
XX  
XX US2003109455-A1.  
XX  
XX 12-JUN-2003.  
XX  
XX 30-OCT-2002; 2002US-00283550.  
XX  
XX 31-OCT-2001; 2001US-0335017P.  
XX  
XX (HAVI/) HAVIV F.  
XX (BRAD/) BRADLEY M F.  
XX  
XX Haviv F, Bradley MF;  
XX WPI; 2003-843101/78.  
XX  
XX New hepta-, octa-, or nonapeptide compounds useful for inhibiting  
XX angiogenesis, treating cancer in mammal, or for prevention of other  
XX diseases such as autoimmune diseases.  
XX  
XX Claim 13; Page 25; 26pp; English.  
XX  
XX The invention describes a hepta-, octa-, or nonapeptide compound (I). (I)  
XX has antiangiogenic, cytostatic, immunosuppressive, antirheumatic, and  
XX antiarthritic activities. (I) or a salt of it, is used in a  
XX pharmaceutical composition for inhibiting angiogenesis or for treating  
XX cancer in a mammal. (I) Can be also used for the treatment or prevention  
XX of other diseases, such as autoimmune diseases, e.g. rheumatoid, immune  
XX and degenerative arthritis. This is the amino acid sequence of an  
XX angiogenesis inhibiting peptide of the invention.  
XX Sequence 8 AA;  
SQ

CC angiogenesis inhibiting peptide of the invention.

XX SQ Sequence 8 AA;

Query Match 78.1%; Score 25; DB 7; Length 8;

Best Local Similarity 75.0%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 2;

QY 3 GVXXIRP 10

Db 1 GVXSTIRP 8

RESULT 4

AD83634

ID ADE83634 standard; peptide; 8 AA.

XX AC ADE83634;

DT 29-JAN-2004 (first entry)

XX DE Antiangiogenic peptide.

XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vulnerary;  
 KW antiinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "acetylated"

FT Modified-site 3 /note= "D-form residue alioisoleucyl"

FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"

FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"

XX WO2003037268-A2.

XX PD 08-MAY-2003.

XX PF 30-OCT-2002; 2002WO-US034811.

XX PR 31-OCT-2001; 2001US-00000681.

XX PR 04-OCT-2002; 2002US-00263812.

XX XX (ABBO ) ABBOTT LAB.

XX PA Haviv F, Bradley MF;

XX PI WPI; 2003-617886/58.

XX DE New hepta-, octa- and nona-peptide compounds used for treating e.g.

XX FT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic

XX FT retinopathy.

XX XX Claim 13; Page 46; 51pp; English.

XX XX The present invention describes hepta-, octa- and nona-peptide compounds

XX CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,

XX CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,

XX CC antidiabetic, dermatological, antipsoriatic, antiarteriosclerotic,

XX CC cardiant, vulnerary, antiinflammatory, antiulcer and haemostatic

XX CC activities. (II) can be used for treating cancer including primary and

XX CC metastatic solid tumours. (I) can also be used for treating autoimmune

XX CC diseases such as rheumatoid, immune and degenerative arthritis, ocular

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.

XX SQ Sequence 8 AA;

Query Match 78.1%; Score 25; DB 7; Length 8;

Best Local Similarity 75.0%; Pred. No. 1.4e+06; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 2;

QY 3 GVXXIRP 10

Db 1 GVXSSIRP 8

RESULT 5

AD83636

ID ADE83636 standard; peptide; 8 AA.

XX AC ADE83636;

XX DT 29-JAN-2004 (first entry)

XX DE Antiangiogenic peptide.

XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vulnerary;  
 KW antiinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "acetylated"

FT Modified-site 3 /note= "D-form residue alioisoleucyl"

FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"

FT Modified-site 8 /note= "Pro is C-terminally modified with -NHCH2CH3"

XX WO2003037268-A2.

XX PD 08-MAY-2003.

XX PF 30-OCT-2002; 2002WO-US034811.

XX PR 31-OCT-2001; 2001US-00000681.

XX PR 04-OCT-2002; 2002US-00263812.

XX XX (ABBO ) ABBOTT LAB.

XX PA Haviv F, Bradley MF;

XX PI WPI; 2003-617886/58.

XX DE New hepta-, octa- and nona-peptide compounds used for treating e.g.

XX FT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic

XX FT retinopathy.

XX XX Claim 13; Page 46; 51pp; English.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antiarteriosclerotic,  
 CC cardiant, vulnary, antiinflammatory, antiulcer and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumours. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 78.1%; Score 25; DB 7; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 |||||  
 Db 1 GVXSTIRP 8

RESULT 6  
 ADE83635  
 ID ADE83635 standard; peptide; 8 AA.

XX ADE83635;

DT 29-JAN-2004 (first entry)

DE Antiangiogenic peptide.

XX antiangiogenic; cytostatic; anti-HIV; immunosuppressive; antirheumatic;  
 KW antiarthritic; ophthalmological; antidiabetic; dermatological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vulnary;  
 KW antiinflammatory; antiulcer; haemostatic; cancer; solid tumour;  
 KW autoimmune disease; rheumatoid arthritis; immune arthritis;  
 KW degenerative arthritis; ocular disease; skin disease;  
 KW blood vessel disease; Osler-Webber Syndrome; myocardial angiogenesis;  
 KW plaque neovascularisation; telangiectasia; haemophilia; angiofibroma;  
 KW wound granulation; birth control; angiogenesis; cat scratch disease;  
 KW ulcer.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "acetylated"  
 FT Modified-site 3  
 FT Modified-site 8 /note= "D-form residue alioisoleucyl"  
 FT Modified-site 9 /note= "Pro is C-terminally modified with -NHCH2CH3"

XX WO2003037269-A2.

XX 08-MAY-2003.

XX 30-OCT-2002; 2002WO-US034811.

XX 31-OCT-2001; 2001US-00000681.

XX 04-OCT-2002; 2002US-00263812.

XX (ABBO ) ABBOTT LAB.

XX

PI Haviv F, Bradley MF;  
 DR WPI; 2003-617886/58.

XX New hepta-, octa- and nona-peptide compounds used for treating e.g.  
 PT cancer, rheumatoid arthritis, psoriasis, scleroderma and diabetic  
 PT retinopathy.

XX Claim 13; Page 46; 5lpp; English.

XX The present invention describes hepta-, octa- and nona-peptide compounds  
 CC (I) having antiangiogenic activity. (I) also have cytostatic, anti-HIV,  
 CC immunosuppressive, antirheumatic, antiarthritic, ophthalmological,  
 CC antidiabetic, dermatological, antipsoriatic, antiarteriosclerotic,  
 CC cardiant, vulnary, antiinflammatory, antiulcer and haemostatic  
 CC activities. (I) can be used for treating cancer including primary and  
 CC metastatic solid tumours. (I) can also be used for treating autoimmune  
 CC diseases such as rheumatoid, immune and degenerative arthritis, ocular  
 CC diseases, skin diseases, blood vessel diseases, Osler-Webber Syndrome,  
 CC myocardial angiogenesis, plaque neovascularisation, telangiectasia,  
 CC haemophilic joints, angiofibroma and wound granulation, and for the  
 CC treatment of diseases involving excessive or abnormal stimulation of  
 CC endothelial cells. (I) can also be used as birth control agents and for  
 CC the treatment of diseases involving angiogenesis as a pathologic  
 CC consequence such as cat scratch disease and ulcers, and to reduce  
 CC bleeding by administration prior to surgery such as for the treatment of  
 CC tumours. The present sequence represents an antiangiogenic peptide, which  
 CC is used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 78.1%; Score 25; DB 7; Length 8;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 |||||  
 Db 1 GVXSTIRP 8

RESULT 7

AAV67983

ID AAV67983 standard; peptide; 9 AA.

XX AAV67983;

DT 11-APR-2000 (first entry)

DE Antiangiogenic peptide #12.

XX antiangiogenic; antiarthritic; cytostatic; antipsoriatic; arthritis;  
 KW angiogenesis; inhibitor; cancer; psoriasis; infection; ocular disease;  
 KW macular degeneration; diabetic retinopathy; tumour metastasis;  
 KW autoimmune disease; neovascularisation; Crohn's disease; birth control;  
 KW cat scratch disease.

XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1 /label= MeCly  
 FT /note= "Sarcosine (methylglycine); acetylated"  
 FT Misc-difference 4  
 FT /note= "D form residue"  
 FT Modified-site 9  
 FT /note= "Pro is modified to ProNHCH2CH3"

XX WO9961476-A1.

XX 02-DEC-1999.

XX 21-MAY-1999; 99WO-US011448.

XX



PR 22-MAY-1998; 98US-00083745.  
 PR 16-FEB-1999; 99US-00250574.  
 PR 26-MAR-1999; 99US-00277465.  
 XX (ABSO ) ABBOTT LAB.  
 PA  
 XX Henkin J, Haviv F, Bradley MF, Kalvin DM, Schneider AJ;  
 PI WPI; 2000-072606/06.  
 XX  
 DR New anti-angiogenic peptides, used for treating e.g. cancer, arthritis,  
 PT psoriasis, or angiogenesis of the eye associated with infection or  
 PT surgical intervention, macular degeneration and diabetic retinopathy.  
 PT  
 XX Claim 12; Page 75; 223pp; English.  
 PS  
 XX AAY67972 to AAY67999 represent novel anti-angiogenic peptides. Peptide  
 CC from the present invention have the formula: A0-A1-A2-A3-A4-A5-A6-A7-A8-  
 CC A9-A10 where A0 is selected from hydrogen or an acyl group; A10 is a  
 CC hydroxyl group or an amino acid amide; and A1 to A9 are amino acyl  
 CC residues. The peptides are used for anti-angiogenesis therapy in  
 CC patients. They are used to treat e.g. cancer, arthritis, psoriasis,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention, macular degeneration and diabetic retinopathy. They can  
 CC prevent tumour metastases. Further uses include treatment and prophylaxis  
 CC of autoimmune diseases, various ocular diseases e.g. diabetic  
 CC retinopathy, and other abnormal neovascularisation conditions of the eye,  
 CC skin diseases e.g. psoriasis, diseases characterised by excessive or  
 CC abnormal stimulation of endothelial cells, e.g. Crohn's disease. They can  
 CC also be used as a birth control agent, inhibiting ovulation and placental  
 CC establishment, and to treat diseases that have angiogenesis as a  
 CC pathological consequence e.g. cat scratch disease and are useful to  
 CC reduce bleeding by administration prior to surgery, especially for the  
 CC treatment of tumours. They can also be used to isolate a receptor from an  
 CC endothelial cell  
 CC  
 XX Sequence 9 AA;  
 XX  
 SQ  
 Query Match 78.1%; Score 25; DB 3; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 DB 2 GVTAIRP 9  
 RESULT 8  
 AAUI5042  
 ID AAUI5042 standard; peptide; 9 AA.  
 AC  
 XX AAUI5042;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Antiangiogenic peptide #69 useful for inhibiting angiogenesis.  
 XX  
 KW Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytosstatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT  
 FT Modified-site 4  
 FT /note= "D-form residue, alloisoleucine"  
 FT Modified-site 6  
 FT /note= "N-Methyl serine"  
 FT

FT Modified-site 9 /note= "Modified by NH-ethyl"  
 FT  
 XX WO200138397-A1.  
 PN  
 XX 31-MAY-2001.  
 PD  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 PF  
 XX  
 XX 22-NOV-1999; 99US-00447099.  
 PR  
 XX 31-OCT-2000; 2000US-00702649.  
 PR  
 XX (ABSO ) ABBOTT LAB.  
 PA  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 PI WPI; 2001-521804/57.  
 XX  
 DR New N-alkylated peptide compounds useful for treating e.g. cancer,  
 XX autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 XX diabetic retinopathy.  
 PT  
 XX Claim 39; Page 89; 95pp; English.  
 PS  
 XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells including hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC scleroderma and hypertrrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #69  
 XX Sequence 9 AA;  
 XX  
 SQ  
 Query Match 78.1%; Score 25; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 DB 2 GVXTSIRP 9  
 RESULT 9  
 AAUI5043  
 ID AAUI5043 standard; peptide; 9 AA.  
 AC  
 XX AAUI5043;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Antiangiogenic peptide #70 useful for inhibiting angiogenesis.  
 XX  
 KW Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytosstatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.  
 XX

```

FH Key Location/Qualifiers
FT Modified-site 1
FT /label= MeGly
FT /note= "N-methyl-glycine (sarcosine), additionally
FT modified by N-terminal acetyl"
FT Misc-difference 4
FT /note= "D-form residue"
FT Modified-site 6
FT /note= "N-methyl serine"
FT Modified-site 9
FT /note= "Modified by NH-ethyl"
FT XX
FT XX
PN WO200138397-A1.
PD 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032105.
XX
XX 22-NOV-1999; 99US-00447099.
PR 31-OCT-2000; 2000US-00702649.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;
XX WPI; 2001-521804/57.
XX
XX New N-alkylated peptide compounds useful for treating e.g. cancer,
XX autoimmune diseases, arthritis, psoriasis, macular degeneration and
XX diabetic retinopathy.
XX
XX Claim 39; Page 89; 95pp; English.
XX
XX The present invention relates to novel synthetic antiangiogenic peptides
XX (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The
XX antiangiogenic peptides are useful for isolating a receptor from an
XX endothelial cell. The peptides of the invention are also useful for
XX treating cancer, arthritis, psoriasis and other skin diseases,
XX angiogenesis of the eye associated with infection or surgical
XX intervention and other ocular diseases, cat scratch disease, ulcers,
XX macular degeneration and diabetic retinopathy, blood vessel diseases e.g.
XX haemangiomas and capillary action within atherosclerotic plaques,
XX autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,
XX Osler-Webber syndrome, myocardial angiogenesis, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
XX wound granulation, excessive or abnormal stimulation of endothelial
XX cells, including intestinal adhesions, Crohn's disease, atherosclerosis,
XX scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic
XX peptides are also useful as birth control agents by inhibiting ovulation
XX and to reduce bleeding by administration before surgery. The peptides of
XX the invention exhibit improved metabolic stability, improved oral
XX pharmacokinetics, increased water solubility, and improved oral
XX availability. The present sequence represents antiangiogenic peptide #70
XX
SQ Sequence 9 AA;
Query Match 78.1%; Score 25; DB 4; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GVXXXIRP 10
DB 2 GVITGIRP 9
RESULT 10
AAU15044
ID AAU15044 standard; peptide; 9 AA.
XX
XX AAU15044;
XX
XX 04-DEC-2001 (first entry)
XX

```

```

DE Antiangiogenic peptide #71 useful for inhibiting angiogenesis.
XX
XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;
XX arthritis; skin disease; ocular disease; diabetic retinopathy;
XX blood vessel disease; atherosclerosis; autoimmune disease; birth control;
XX cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Modified-site 1
FT /label= MeGly
FT /note= "N-methyl-glycine (sarcosine), additionally
FT modified by N-terminal acetyl"
FT
FT Modified-site 4
FT /note= "D-form residue, alioisoleucine"
FT
FT Modified-site 5
FT /note= "N-Methyl serine"
FT
FT Modified-site 9
FT /note= "Modified by NH-ethyl"
XX
XX WO200138397-A1.
XX
XX 31-MAY-2001.
XX
XX 22-NOV-2000; 2000WO-US032105.
XX
XX 22-NOV-1999; 99US-00447099.
PR 31-OCT-2000; 2000US-00702649.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Haviv F, Henkin J, Bradley MF, Kalvin DM;
XX WPI; 2001-521804/57.
XX
XX New N-alkylated peptide compounds useful for treating e.g. cancer,
XX autoimmune diseases, arthritis, psoriasis, macular degeneration and
XX diabetic retinopathy.
XX
XX Claim 39; Page 89; 95pp; English.
XX
XX The present invention relates to novel synthetic antiangiogenic peptides
XX (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The
XX antiangiogenic peptides are useful for isolating a receptor from an
XX endothelial cell. The peptides of the invention are also useful for
XX treating cancer, arthritis, psoriasis and other skin diseases,
XX angiogenesis of the eye associated with infection or surgical
XX intervention and other ocular diseases, cat scratch disease, ulcers,
XX macular degeneration and diabetic retinopathy, blood vessel diseases e.g.
XX haemangiomas and capillary action within atherosclerotic plaques,
XX autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,
XX Osler-Webber syndrome, myocardial angiogenesis, plaque
XX neovascularisation, telangiectasia, haemophilic joints, angiofibroma,
XX wound granulation, excessive or abnormal stimulation of endothelial
XX cells, including intestinal adhesions, Crohn's disease, atherosclerosis,
XX scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic
XX peptides are also useful as birth control agents by inhibiting ovulation
XX and to reduce bleeding by administration before surgery. The peptides of
XX the invention exhibit improved metabolic stability, improved oral
XX pharmacokinetics, increased water solubility, and improved oral
XX availability. The present sequence represents antiangiogenic peptide #71
XX
SQ Sequence 9 AA;
Query Match 78.1%; Score 25; DB 4; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.4e+06;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 GVXXXIRP 10
DB 2 GVXSIRP 9

```

## RESULT 11

AAU15027  
 ID AAU15027 standard; peptide; 9 AA.  
 XX  
 AC AAU15027;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Antiangiogenic peptide #54 useful for inhibiting angiogenesis.  
 XX  
 KW Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Misc-difference 4  
 FT /note= "D-form residue"  
 FT Modified-site 6  
 FT /note= "N-Methyl alanine"  
 FT Modified-site 9  
 FT /note= "Modified by NH-ethyl"  
 XX  
 FN WO200138397-A1.  
 XX  
 PD 31-MAY-2001.  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 PF  
 XX 22-NOV-1999; 99US-00447099.  
 PR  
 PR 31-OCT-2000; 2000US-00702649.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.  
 DR  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 PT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 PT diabetic retinopathy.  
 PS  
 PS Claim 39; Page 89; 95pp; English.  
 XX  
 CC The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques.  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #54  
 XX Sequence 9 AA;

Query Match 78.1%; Score 25; DB 4; Length 9;  
 Best Local Similarity 62.5%; Pred. No. 1.4e+06;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 DB 2 GVITAIRP 9  
 RESULT 12  
 AAU15041  
 ID AAU15041 standard; peptide; 9 AA.  
 XX  
 AC AAU15041;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Antiangiogenic peptide #68 useful for inhibiting angiogenesis.  
 XX  
 KW Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cytostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Modified-site 4  
 FT /note= "D-form residue, allosioleucine"  
 FT Modified-site 6  
 FT /note= "N-Methyl serine"  
 FT Modified-site 9  
 FT /note= "Modified by NH-ethyl"  
 XX  
 FN WO200138397-A1.  
 XX  
 PD 31-MAY-2001.  
 XX  
 XX 22-NOV-2000; 2000WO-US032105.  
 PF  
 XX 22-NOV-1999; 99US-00447099.  
 PR  
 PR 31-OCT-2000; 2000US-00702649.  
 XX  
 XX (ABBO ) ABBOTT LAB.  
 XX  
 XX Haviv F, Henkin J, Bradley MF, Kalvin DM;  
 XX WPI; 2001-521804/57.  
 DR  
 XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 PT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 PT diabetic retinopathy.  
 PS  
 PS Claim 39; Page 89; 95pp; English.  
 XX  
 CC The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for  
 CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques.  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #54  
 XX Sequence 9 AA;

CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved  
 CC pharmacokinetics, increased water solubility, and improved oral  
 CC availability. The present sequence represents antiangiogenic peptide #68  
 XX  
 SQ Sequence 9 AA;

Query Match 78.1%; Score 25; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 ||| |||  
 Db 2 GVXSIRP 9

RESULT 13  
 AAU15016  
 ID AAU15016 standard; peptide; 9 AA.  
 XX  
 AC AAU15016;

DT 04-DEC-2001 (first entry)  
 XX

DE Antiangiogenic peptide #43 useful for inhibiting angiogenesis.

XX Antiangiogenic; angiogenesis inhibitor; endothelial cell; cancer;  
 KW arthritis; skin disease; ocular disease; diabetic retinopathy;  
 KW blood vessel disease; atherosclerosis; autoimmune disease; birth control;  
 KW cyostatic; anti arthritic; anti rheumatic; anti ulcer; anti psoriatic.  
 OS  
 XX Synthetic.

XX Key Location/Qualifiers  
 FH Modified-site 1 /label= MeGly  
 FT /note= "N-methyl-glycine (sarcosine), additionally  
 FT modified by N-terminal acetyl"  
 FT Modified-site 4  
 FT /note= "D-form residue, alloisoleucine"  
 FT Modified-site 5  
 FT /note= "N-Methyl serine"  
 FT Modified-site 9  
 FT /note= "Modified by NH-ethyl"  
 FT  
 XX WO200138397-A1.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032105.

XX 22-NOV-1999; 99US-00447099.

PR 31-OCT-2000; 2000US-00702649.

XX (ABBO ) ABBOTT LAB.

XX Haviv F, Henkin J, Bradley MF, Kalvin DM;

XX WPI; 2001-521804/57.

XX New N-alkylated peptide compounds useful for treating e.g. cancer,  
 FT autoimmune diseases, arthritis, psoriasis, macular degeneration and  
 FT diabetic retinopathy.

XX Claim 38, Page 88; 95pp; English.

XX The present invention relates to novel synthetic antiangiogenic peptides  
 CC (AAU14974-AAU15052) which can act as angiogenesis inhibitors. The  
 CC antiangiogenic peptides are useful for isolating a receptor from an  
 CC endothelial cell. The peptides of the invention are also useful for

CC treating cancer, arthritis, psoriasis and other skin diseases,  
 CC angiogenesis of the eye associated with infection or surgical  
 CC intervention and other ocular diseases, cat scratch disease, ulcers,  
 CC macular degeneration and diabetic retinopathy, blood vessel diseases e.g.  
 CC haemangiomas and capillary action within atherosclerotic plaques,  
 CC autoimmune diseases e.g. rheumatoid, immune and degenerative arthritis,  
 CC Osler-Weber syndrome, myocardial angiogenesis, plaque  
 CC neovascularisation, telangiectasia, haemophilic joint, angiodioma,  
 CC wound granulation, excessive or abnormal stimulation of endothelial  
 CC cells, including intestinal adhesions, Crohn's disease, atherosclerosis,  
 CC scleroderma and hypertrophic scars, i.e. keloids. The antiangiogenic  
 CC peptides are also useful as birth control agents by inhibiting ovulation  
 CC and to reduce bleeding by administration before surgery. The peptides of  
 CC the invention exhibit improved metabolic stability, improved oral  
 CC pharmacokinetics, increased water solubility, and improved oral  
 XX availability. The present sequence represents antiangiogenic peptide #43  
 XX  
 SQ Sequence 9 AA;

Query Match 78.1%; Score 25; DB 4; Length 9;  
 Best Local Similarity 75.0%; Pred. No. 1.4e+06;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 ||| |||  
 Db 2 GVXSIRP 9

RESULT 14

AAO26627

ID AAO26627 standard; peptide; 9 AA.

XX AAO26627;

XX 28-MAR-2003 (first entry)

XX Anti-angiogenic peptide #10.

XX Cytostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;  
 KW dermatological; immunosuppressive; cardiant; vulnerary; antiulcer;  
 KW antiarteriosclerotic; angiogenesis inhibitor; cancer; arthritis;  
 KW psoriasis; angiogenesis; eye; infection; surgical intervention;  
 KW macular degeneration; diabetic retinopathy; autoimmune disease;  
 KW ocular disease; skin disease; blood vessel disease; telangiectasia;  
 KW Osler-Weber Syndrome; myocardial angiogenesis; haemophilic joint;  
 KW plaque neovascularisation; angiodioma; wound granulation; scleroderma;  
 KW atherosclerosis; intestinal adhesion; Crohn's disease; hypertrophic scar;  
 KW birth control agent; cat scratch disease; ulcer; angiogenic.  
 XX Synthetic.

XX Key Location/Qualifiers

FH Modified-site 1 /label= MeGly

FT /note= "This sarcosine residue is modified by N-Ac"

FT Modified-site 4 /note= "This residue is modified to become allole"

FT Misc-difference 4 /note= "This is a D-form residue"

FT Modified-site 9 /note= "Residue is modified by NHCH2CH3"

XX WO200283065-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US011027.

XX 11-APR-2001; 2001US-00832733.

XX (ABBO ) ABBOTT LAB.

XX Henkin J, Haviv F, Bradley MF, Douglas KM, Schneider AJ;

```

XX WPI; 2003-111805/10.
XX
XX New peptides are angiogenesis inhibitors used for treating e.g. cancer,
XX arthritis and psoriasis.
XX
XX Claim 12; Page 31; 33pp; English.
XX
XX The invention relates to novel peptides for use as angiogenesis
XX inhibitors. Used as angiogenesis inhibitors used for treating cancer,
XX arthritis, psoriasis, angiogenesis of the eye associated with infection
XX or surgical intervention, macular degeneration and diabetic retinopathy.
XX The novel peptides are also used for treating autoimmune diseases, ocular
XX diseases, skin diseases, blood vessel diseases, Osler Webber Syndrome,
XX myocardial angiogenesis, plaque neovascularisation, telangiectasia,
XX haemophilic joints, angiofibroma, wound granulation, diseases
XX characterised by excessive or abnormal stimulation of endothelial cells
XX including intestinal adhesions, Crohn's disease, atherosclerosis,
XX scleroderma and hypertrophic scars. The novel peptides are also used as
XX birth control agents and for treating cat scratch disease and ulcers.
XX This sequence represents one of the angiogenic peptides of the invention
XX
XX Sequence 9 AA;
XX
Query Match      78.1%; Score 25; DB 6; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
Db 2 GVISTIRP 9

RESULT 15
AAO26633
ID AAO26633 standard; peptide; 9 AA.
AC
AC AAO26633;
XX
XX 28-MAR-2003 (first entry)
XX
XX Anti-angiogenic peptide #16.
XX
XX Cystostatic; antiarthritic; antipsoriatic; antidiabetic; ophthalmological;
XX dermatological; immunosuppressive; cardiac; vulnerable; antiulcer;
XX antiarteriosclerotic; angiogenesis inhibitor; cancer; arthritis;
XX psoriasis; angiogenesis; eye; infection; surgical intervention;
XX macular degeneration; diabetic retinopathy; autoimmune disease;
XX ocular disease; skin disease; blood vessel disease; telangiectasia;
XX Osler Webber Syndrome; myocardial angiogenesis; haemophilic joint;
XX plaque neovascularisation; angiofibroma; wound granulation; scleroderma;
XX atherosclerosis; intestinal adhesion; Crohn's disease; hypertrophic scar;
XX birth control agent; cat scratch disease; ulcer; angiogenic.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /label= MeGly
XX FT /note= "This sarcosine residue is modified by N-Ac"
XX FT Misc-difference 4 /note= "This is a D-form residue"
XX FT Modified-site 5 /note= "Residue is modified to become allothr"
XX FT Modified-site 9 /note= "Residue is modified by NHCH2CH3"
XX
XX WO200283065-A2.
XX
XX 24-OCT-2002.
XX
XX 10-APR-2002; 2002WO-US011027.
XX

```

```

PR 11-APR-2001; 2001US-00832733.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Henkin J, Haviv F, Bradley MF, Douglas KM, Schneider AJ;
XX
XX WPI; 2003-111805/10.
XX
XX New peptides are angiogenesis inhibitors used for treating e.g. cancer,
XX arthritis and psoriasis.
XX
XX Claim 12; Page 31; 33pp; English.
XX
XX The invention relates to novel peptides for use as angiogenesis
XX inhibitors. Used as angiogenesis inhibitors used for treating cancer,
XX arthritis, psoriasis, angiogenesis of the eye associated with infection
XX or surgical intervention, macular degeneration and diabetic retinopathy.
XX The novel peptides are also used for treating autoimmune diseases, ocular
XX diseases, skin diseases, blood vessel diseases, Osler Webber Syndrome,
XX myocardial angiogenesis, plaque neovascularisation, telangiectasia,
XX haemophilic joints, angiofibroma, wound granulation, diseases
XX characterised by excessive or abnormal stimulation of endothelial cells
XX including intestinal adhesions, Crohn's disease, atherosclerosis,
XX scleroderma and hypertrophic scars. The novel peptides are also used as
XX birth control agents and for treating cat scratch disease and ulcers.
XX This sequence represents one of the angiogenic peptides of the invention
XX
XX Sequence 9 AA;
XX
Query Match      78.1%; Score 25; DB 6; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
Db 2 GVISTIRP 9

Search completed: April 1, 2004, 17:35:25
Job time : 53.6842 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:30:38 ; Search time 15.6316 Seconds  
(without alignments)  
36.329 Million cell updates/sec

Title: US-09-833-196-6  
Perfect score: 32  
Sequence: 1 XGVXXIXRP 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	78.1	479	4	US-09-252-991A-23144
2	25	78.1	728	4	US-09-252-991A-31891
3	24	75.0	161	4	US-09-230-196-8
4	24	75.0	176	4	US-09-252-991A-30148
5	24	75.0	176	4	US-09-489-039A-11985
6	24	75.0	208	4	US-09-252-991A-23734
7	24	75.0	240	2	US-08-114-555A-6
8	24	75.0	240	2	US-08-114-555A-8
9	24	75.0	240	3	US-08-559-397A-12
10	24	75.0	240	3	US-08-559-397A-14
11	24	75.0	252	4	US-09-328-352-8187
12	24	75.0	330	4	US-09-489-039A-7373
13	24	75.0	395	4	US-09-489-039A-9276
14	24	75.0	475	4	US-09-252-991A-32806
15	24	74.8	748	4	US-09-252-991A-21696
16	23	71.9	18	4	US-08-442-001C-82
17	23	71.9	40	2	US-08-919-724-5
18	23	71.9	76	4	US-09-252-991A-25888
19	23	71.9	100	4	US-09-543-681A-5475
20	23	71.9	140	4	US-09-540-236-2611
21	23	71.9	153	4	US-09-252-991A-32200
22	23	71.9	182	4	US-09-198-452A-846
23	23	71.9	187	4	US-09-489-039A-13324
24	23	71.9	252	4	US-09-107-532A-4457
25	23	71.9	271	4	US-09-252-991A-26235
26	23	71.9	294	4	US-09-252-991A-29464
27	23	71.9	323	4	US-09-489-039A-7409

Sequence 25209, A  
Sequence 13707, A  
Sequence 29467, A  
Sequence 72, Appl  
Sequence 27787, A  
Sequence 6748, Ap  
Sequence 3023, Ap  
Sequence 8036, Ap  
Sequence 25823, A  
Sequence 1, Appli  
Sequence 5, Appli  
Sequence 17231, A  
Sequence 30858, A  
Sequence 27028, A  
Sequence 13699, A  
Sequence 279, App  
Sequence 284, App

28 71.9 352 4 US-09-252-991A-25209  
29 71.9 498 4 US-09-489-039A-13707  
30 71.9 510 4 US-09-252-991A-29467  
31 71.9 547 4 US-10-164-595-72  
32 71.9 551 4 US-09-252-991A-27787  
33 71.9 661 4 US-09-328-352-8036  
34 71.9 727 4 US-09-540-236-3023  
35 71.9 729 4 US-09-328-352-8036  
36 71.9 792 4 US-09-252-991A-25823  
37 71.9 1066 2 US-08-633-770A-1  
38 71.9 1066 4 US-09-280-197-5  
39 71.9 2294 4 US-09-252-991A-17231  
40 68.8 136 4 US-09-252-991A-30858  
41 68.8 160 4 US-09-252-991A-27028  
42 68.8 187 4 US-09-252-991A-27028  
43 68.8 187 4 US-09-489-039A-13699  
44 68.8 240 4 US-09-072-596-279  
45 68.8 240 4 US-09-072-967-284

## ALIGNMENTS

RESULT 1  
US-09-252-991A-23144  
; Sequence 23144, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23144  
; LENGTH: 479  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23144

Query Match 78.1%; Score 25; DB 4; Length 479;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 3 GYXXIXRP 10  
Db 219 GVSSAVRP 226

RESULT 2  
US-09-252-991A-31891  
; Sequence 31891, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31891  
; LENGTH: 728  
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31891

Query Match 78.1%; Score 25; DB 4; Length 728;  
Best Local Similarity 50.0%; Pred. No. 1.1e-02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|||  
Db 474 GVASVRP 481

RESULT 3  
US-09-230-196-8  
; Sequence 8, Application US/09230196  
; Patent No. 6307035  
; GENERAL INFORMATION:  
; APPLICANT: Rauscher III, Frank J.  
; APPLICANT: Jensen, David E.  
; TITLE OF INVENTION: BRCA Associated Protein (BAP-1) and  
; TITLE OF INVENTION: Uses Therefor  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr., PO Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/230,196  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/022,997  
; FILING DATE: 02-AUG-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/038,109  
; FILING DATE: 19-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: WST68BUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 161 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-230-196-8

Query Match 75.0%; Score 24; DB 4; Length 161;  
Best Local Similarity 62.5%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|||  
Db 8 GVRAQIRP 15

RESULT 4  
US-09-252-991A-30148  
; Sequence 30148, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30148  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30148

Query Match 75.0%; Score 24; DB 4; Length 176;  
Best Local Similarity 50.0%; Pred. No. 46;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|||  
Db 45 GVTARVP 52

RESULT 5  
US-09-489-039A-11985  
; Sequence 11985, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Bretton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11985  
; LENGTH: 176  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11985

Query Match 75.0%; Score 24; DB 4; Length 176;  
Best Local Similarity 62.5%; Pred. No. 46;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|||  
Db 150 GVTGVRP 157

RESULT 6  
US-09-252-991A-23734  
; Sequence 23734, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 23734

LENGTH: 208  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-23734

Query Match 75.0%; Score 24; DB 4; Length 208;  
Best Local Similarity 50.0%; Pred. No. 55;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|| :||  
Db 5 GVASEVRP 12

RESULT 7  
US-08-114-555A-6  
Sequence 6, Application US/08114555A  
Patent No. 5854392  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.  
APPLICANT: Kozlowski, Michael L.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,555A  
FILING DATE: 30-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-115  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-114-555A-6

Query Match 75.0%; Score 24; DB 2; Length 240;  
Best Local Similarity 50.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|| :||  
Db 3 GVTTSRNP 10

RESULT 8  
US-08-114-555A-8  
Sequence 8, Application US/08114555A  
Patent No. 5854392  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.

APPLICANT: Kozlowski, Michael R.  
APPLICANT: Neve, Rachael L.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF beta-APP-C100  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/114,555A  
FILING DATE: 30-AUG-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-115  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-114-555A-8

Query Match 75.0%; Score 24; DB 2; Length 240;  
Best Local Similarity 50.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|| :||  
Db 3 GVTTSRNP 10

RESULT 9  
US-08-559-397A-12  
Sequence 12, Application US/08559397A  
Patent No. 6083713  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.  
APPLICANT: Kozlowski, Michael R.  
TITLE OF INVENTION: CLONING AND EXPRESSION OF  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,397A  
FILING DATE: 15-NOV-1995



CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-559-397A-12

Query Match 75.0%; Score 24; DB 3; Length 240;  
Best Local Similarity 50.0%; Pred. No. 64;  
Matches 4; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 3 GVXXXIRP 10  
DB 3 GVTTSIRP 10

RESULT 10  
US-08-559-397A-14  
Sequence 14, Application US/08559397A  
Patent No. 6083713  
GENERAL INFORMATION:  
APPLICANT: Manly, Susan P.  
APPLICANT: Kozlowski, Michael R.  
APPLICANT: Neve, Rachael L.

TITLE OF INVENTION: CLONING AND EXPRESSION OF  
TITLE OF INVENTION: BETA APP-C100 RECEPTOR (C100-R)  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036/2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,397A  
FILING DATE: 15-NOV-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 6013-135  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-559-397A-14

Query Match 75.0%; Score 24; DB 3; Length 240;  
Best Local Similarity 50.0%; Pred. No. 64;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
DB 3 GVTTSIRP 10

RESULT 11  
US-09-328-352-8187  
Sequence 8187, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC99-039A  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 8187  
LENGTH: 252  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-8187

Query Match 75.0%; Score 24; DB 4; Length 252;  
Best Local Similarity 50.0%; Pred. No. 67;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
DB 74 GVAAVRP 81

RESULT 12  
US-09-489-039A-7373  
Sequence 7373, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 7373  
LENGTH: 330  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7373

Query Match 75.0%; Score 24; DB 4; Length 330;  
Best Local Similarity 62.5%; Pred. No. 88;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
DB 42 GVSNIRP 49

RESULT 13  
US-09-489-039A-9276  
Sequence 9276, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A

Best Local Similarity 62.5%; Pred. No. 2e+02; Indels 3; Gaps 0;  
 Matches 5; Conservative 0; Mismatches 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 Db 740 GVMTHIRP 747  
 Search completed: April 1, 2004, 17:42:13  
 Job time : 16.6316 secs

CURRENT FILING DATE: 2000-01-27  
 PRIOR APPLICATION NUMBER: US 60/117,747  
 PRIOR FILING DATE: 1999-01-29  
 NUMBER OF SEQ ID NOS: 14342  
 SEQ ID NO 9276  
 LENGTH: 395  
 TYPE: PRT  
 ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-9276  
 Query Match 75.0%; Score 24; DB 4; Length 395;  
 Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 Db 140 GIEAIRP 147  
 RESULT 14  
 US-09-252-991A-32806  
 Sequence 32806, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 32806  
 LENGTH: 475  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-32806

Query Match 75.0%; Score 24; DB 4; Length 475;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 Db 49 GVLGIRP 56  
 RESULT 15  
 US-09-252-991A-21696  
 Sequence 21696, Application US/09252991A  
 Patent No. 6551795  
 GENERAL INFORMATION:  
 APPLICANT: Marc J. Rubenfield et al.  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 107196.136  
 CURRENT APPLICATION NUMBER: US/09/252,991A  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US 60/074,788  
 PRIOR FILING DATE: 1998-02-18  
 PRIOR APPLICATION NUMBER: US 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 21696  
 LENGTH: 748  
 TYPE: PRT  
 ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-21696  
 Query Match 75.0%; Score 24; DB 4; Length 748;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:35:28 ; Search time 37.4386 Seconds

(without alignments)  
77.074 Million cell updates/sec

Title: US-09-833-196-6  
Perfect score: 32  
Sequence: 1 XXGVXXIRPX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1069545 seqs, 262320428 residues  
Total number of hits satisfying chosen parameters: 1069545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	25	78.1	89	12	US-10-424-599-188130
2	25	78.1	139	14	US-10-187-267A-17
3	25	78.1	365	12	US-10-425-114-47548
4	25	78.1	482	15	US-10-369-493-17095
5	25	78.1	2796	9	US-09-820-759-114
6	25	78.1	2796	10	US-09-751-708A-114
7	24	75.0	133	15	US-10-121-016-7
8	24	75.0	133	15	US-10-121-016-69
9	24	75.0	139	12	US-10-425-114-51464
10	24	75.0	143	12	US-10-424-599-271715
11	24	75.0	158	12	US-10-424-599-276515
12	24	75.0	204	14	US-10-029-386-32508
13	24	75.0	312	9	US-10-156-761-11045
14	24	75.0	314	9	US-09-738-626-6809
15	24	75.0	386	12	US-10-424-599-269274

Sequence 276516, A  
Sequence 19542, A  
Sequence 210, Appl  
Sequence 84, Appl  
Sequence 3577, Ap  
Sequence 1850, Ap  
Sequence 5100, Ap  
Sequence 205572, A  
Sequence 279855, A  
Sequence 154588, A  
Sequence 45189, A  
Sequence 207352, A  
Sequence 45558, A  
Sequence 250928, A  
Sequence 66811, A  
Sequence 145205, A  
Sequence 846, Appl  
Sequence 12794, A  
Sequence 172616, A  
Sequence 278337, A  
Sequence 37632, A  
Sequence 32561, A  
Sequence 14738, A  
Sequence 6611, Appl  
Sequence 4, Appl  
Sequence 265468, A  
Sequence 37945, A  
Sequence 14341, A  
Sequence 14682, A  
Sequence 2875, Ap

US-10-424-599-276516  
US-10-369-493-19542  
US-10-114-270-210  
US-10-093-463-84  
US-10-369-493-3577  
US-10-276-774-1850  
US-09-764-891-5100  
US-10-424-599-205572  
US-10-424-599-279855  
US-10-424-599-154588  
US-09-864-761-45189  
US-10-424-599-207352  
US-09-864-761-46558  
US-10-424-599-250928  
US-10-425-114-66811  
US-10-289-762-846  
US-10-156-761-12794  
US-10-424-599-278337  
US-10-425-114-37632  
US-10-029-386-32561  
US-10-156-761-14738  
US-09-738-626-6611  
US-10-067-482-4  
US-10-424-599-265468  
US-09-864-761-37945  
US-10-369-493-14341  
US-10-156-761-14682  
US-10-364-049-2875

## ALIGNMENTS

RESULT 1  
US-10-424-599-188130  
; Sequence 188130, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-2153223B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 188130  
; LENGTH: 89  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_140896C.1.pep  
US-10-424-599-188130

Query Match 78.1%; Score 25; DB 12; Length 89;  
Best Local Similarity 62.5%; Pred No. 54;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXIRP 10  
|||  
DB 78 GVITSIRP 85

RESULT 2  
US-10-187-267A-17  
; Sequence 17, Application US/10187267A  
; Publication No. US20030124679A1  
; GENERAL INFORMATION:  
; APPLICANT: Short, Jay M.

; APPLICANT: Paradkar, Ashish  
 ; APPLICANT: Varoglu, Mustafa  
 ; APPLICANT: Mathur, Eric J.  
 ; TITLE OF INVENTION: GLYCOSYLATED KINAMYCINS AND METHODS OF  
 ; TITLE OF INVENTION: MAKING AND USING THEM  
 ; FILE REFERENCE: 09010-280001  
 ; CURRENT APPLICATION NUMBER: US/10/187,267A  
 ; CURRENT FILING DATE: 2003-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/301,401  
 ; PRIOR FILING DATE: 2001-06-27  
 ; NUMBER OF SEQ ID NOS: 65  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 17  
 ; LENGTH: 139  
 ; TYPE: PRT  
 ; ORGANISM: Streptomyces murayamaensis ATCC 21414  
 US-10-187-267A-17

Query Match 78.1%; Score 25; DB 14; Length 139;  
 Best Local Similarity 62.5%; Pred. No. 85;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXIRP 10  
 |||||  
 Db 28 GVLTIIRP 35

RESULT 3  
 US-10-425-114-47548  
 ; Sequence 47548, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E.  
 ; APPLICANT: Tabaska, Jack E.  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(5313) B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 47548  
 ; LENGTH: 365  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700071806\_FLI.pep  
 US-10-425-114-47548

Query Match 78.1%; Score 25; DB 12; Length 365;  
 Best Local Similarity 50.0%; Pred. No. 2,2e+02;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXIRP 10  
 |||||  
 Db 20 GVAASVRP 27

RESULT 4  
 US-10-369-493-17095  
 ; Sequence 17095, Application US/10369493  
 ; Publication No. US20030233675A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Hinkle, Gregory J.  
 ; APPLICANT: Slater, Steven C.  
 ; APPLICANT: Goldman, Barry S.  
 ; APPLICANT: Chen, Xianfeng  
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052) B  
 ; CURRENT APPLICATION NUMBER: US/10/369,493  
 ; CURRENT FILING DATE: 2003-02-28  
 ; PRIOR APPLICATION NUMBER: US 60/360,039  
 ; PRIOR FILING DATE: 2002-02-21  
 ; NUMBER OF SEQ ID NOS: 47374  
 ; SEQ ID NO 17095  
 ; LENGTH: 482  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus halodurans  
 US-10-369-493-17095

Query Match 78.1%; Score 25; DB 15; Length 482;  
 Best Local Similarity 62.5%; Pred. No. 3e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXIRP 10  
 |||||  
 Db 173 GVARSIRP 180

RESULT 5  
 US-09-870-759-114  
 ; Sequence 114, Application US/09870759  
 ; Patent No. US20020177551A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Terman, David S  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
 ; FILE REFERENCE: 870759  
 ; CURRENT APPLICATION NUMBER: US/09/870,759  
 ; CURRENT FILING DATE: 2002-01-14  
 ; PRIOR APPLICATION NUMBER: US 60/208,128  
 ; PRIOR FILING DATE: 2000-05-30  
 ; NUMBER OF SEQ ID NOS: 166  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 114  
 ; LENGTH: 2796  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium bovis  
 US-09-870-759-114

Query Match 78.1%; Score 25; DB 9; Length 2796;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXIRP 10  
 |||||  
 Db 2704 GVHTSIRP 2711

RESULT 6  
 US-09-751-708A-114  
 ; Sequence 114, Application US/09751708A  
 ; Publication No. US20030157113A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Terman, David S  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
 ; FILE REFERENCE: 751708  
 ; CURRENT APPLICATION NUMBER: US/09/751,708A  
 ; CURRENT FILING DATE: 2002-10-15  
 ; PRIOR APPLICATION NUMBER: US 60/173,371  
 ; PRIOR FILING DATE: 1999-12-28  
 ; NUMBER OF SEQ ID NOS: 166  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 114  
 ; LENGTH: 2796  
 ; TYPE: PRT  
 ; ORGANISM: Mycobacterium bovis  
 US-09-751-708A-114

Query Match 78.1%; Score 25; DB 10; Length 2796;  
 Best Local Similarity 62.5%; Pred. No. 1.7e+03;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY      3 GVXXXIRP 10
Db      2704 GVHSIRP 2711

RESULT 7
US-10-121-016-7
; Sequence 7, Application US/10121016
; Publication No. US20040010811A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Wangmao Ge
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE OF INVENTION: ENTITLED 162PLE6 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20077.00
; CURRENT APPLICATION NUMBER: US/10/121.016
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-121-016-7

Query Match      75.0%; Score 24; DB 15; Length 133;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 GVXXXIRP 10
Db      121 GVSHIRP 128

RESULT 8
US-10-121-016-69
; Sequence 69, Application US/10121016
; Publication No. US20040010811A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Wangmao Ge
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; FILE OF INVENTION: ENTITLED 162PLE6 USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 51158-20077.00
; CURRENT APPLICATION NUMBER: US/10/121.016
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/283,112
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: US 60/286,630
; PRIOR FILING DATE: 2001-04-25
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-121-016-69

Query Match      75.0%; Score 24; DB 15; Length 133;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      3 GVXXXIRP 10
Db      121 GVSHIRP 128

RESULT 9
US-10-425-114-51464
; Sequence 51464, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51464
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700894905_FLI.pgp
US-10-425-114-51464

Query Match      75.0%; Score 24; DB 12; Length 139;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 GVXXXIRP 10
Db      74.GVAVTRP 81

RESULT 10
US-10-424-599-271715
; Sequence 271715, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271715
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87378C.1.pgp
US-10-424-599-271715

Query Match      75.0%; Score 24; DB 12; Length 143;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

```
QY      3 GVXXXIRP 10
Db      121 GVSHIRP 128

RESULT 9
US-10-425-114-51464
; Sequence 51464, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 51464
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700894905_FLI.pgp
US-10-425-114-51464

Query Match      75.0%; Score 24; DB 12; Length 139;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      3 GVXXXIRP 10
Db      74.GVAVTRP 81

RESULT 10
US-10-424-599-271715
; Sequence 271715, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 271715
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_87378C.1.pgp
US-10-424-599-271715

Query Match      75.0%; Score 24; DB 12; Length 143;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

QY 3 GVXXXIRP 10  
||| :||  
Db 78 GVAVTVRP 85

## RESULT 11

US-10-424-599-276515  
; Sequence 276515, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 276515

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_91712C.1.pcp

US-10-424-599-276515

Query Match 75.0%; Score 24; DB 12; Length 158;

Best Local Similarity 50.0%; Pred. No. 1.8e+02;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
||| :||  
Db 99 GVSTAVRP 106

## RESULT 12

US-10-029-386-32508

; Sequence 32508, Application US/10029386

; Publication No. US20030194704A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

; FILE REFERENCE: AEOMICA-X-2

; CURRENT APPLICATION NUMBER: US/10/029,386

; CURRENT FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 34288

; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

; SEQ ID NO 32508

; LENGTH: 204

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC011469.3

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.95

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4

; OTHER INFORMATION: SWISSPROT HIT: P21817, EVALUATE 1.00e-104

US-10-029-386-32508

Query Match 75.0%; Score 24; DB 14; Length 204;

Best Local Similarity 50.0%; Pred. No. 2.3e+02;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10

Db 121 GVTSLRP 128  
||| :||

## RESULT 13

US-10-156-761-11045

; Sequence 11045, Application US/10156761

; Publication No. US20030119018A1

; GENERAL INFORMATION:

; APPLICANT: OMURA, SATOSHI

; APPLICANT: IKEDA, HARUO

; APPLICANT: ISHIKAWA, JUN

; APPLICANT: HORIKAWA, HIROSHI

; APPLICANT: SHIBA, TADAYOSHI

; APPLICANT: SAKAKI, YOSHIYUKI

; APPLICANT: HATTORI, MASAHIRA

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-262

; CURRENT APPLICATION NUMBER: US/10/156,761

; CURRENT FILING DATE: 2002-05-29

; PRIOR APPLICATION NUMBER: JP 2001-204089

; PRIOR FILING DATE: 2001-05-30

; PRIOR APPLICATION NUMBER: JP 2001-272697

; PRIOR FILING DATE: 2001-08-02

; NUMBER OF SEQ ID NOS: 15109

; SEQ ID NO 11045

; LENGTH: 312

; TYPE: PRT

; ORGANISM: Streptomyces avermitilis

US-10-156-761-11045

Query Match 75.0%; Score 24; DB 14; Length 312;

Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
||| :||  
Db 71 GVRTTVRP 78

## RESULT 14

US-09-738-626-6809

; Sequence 6809, Application US/09738626

; Publication No. US20020197605A1

; GENERAL INFORMATION:

; APPLICANT: NAKAGAWA, SATOSHI

; APPLICANT: MIZOGUCHI, HIROSHI

; APPLICANT: ANDO, SEIKO

; APPLICANT: HAYASHI, MIKIRO

; APPLICANT: OCHIAI, KEIKO

; APPLICANT: YOKOI, HARUHIKO

; APPLICANT: TATEISHI, NAOKO

; APPLICANT: SENOH, AKIHIRO

; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125

; CURRENT APPLICATION NUMBER: US/09/738,626

; CURRENT FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: JP 99/377484

; PRIOR FILING DATE: 1999-12-16

; PRIOR APPLICATION NUMBER: JP 00/159162

; PRIOR FILING DATE: 2000-04-07

; PRIOR APPLICATION NUMBER: JP 00/280988

; PRIOR FILING DATE: 2000-08-03

; NUMBER OF SEQ ID NOS: 7059

; SOFTWARE: PatentIn ver. 3.0

; SEQ ID NO 6809

; LENGTH: 314

; TYPE: PRT

; ORGANISM: Corynebacterium glutamicum

US-09-738-626-6809

```
Query Match      75.0%; Score 24; DB 9; Length 314;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      3 GVXXXIRP 10
      | : |||
Db      46 GITTFIRP 53

RESULT 15
US-10-424-599-269274
; Sequence 269274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 269274
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(386)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85175C.1.pep
US-10-424-599-269274

Query Match      75.0%; Score 24; DB 12; Length 386;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      3 GVXXXIRP 10
      | : |||
Db      185 GVATVRP 192

Search completed: April 1, 2004, 17:45:40
Job time : 38.4386 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:43 ; Search time 11.9649 Seconds  
(without alignments)  
88.434 Million cell updates/sec

Title: US-09-833-196-6  
Perfect score: 32  
Sequence: 1 XXGVXXXIRPX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:.\*  
1: Piri:.\*  
2: Piri2:.\*  
3: Piri3:.\*  
4: Piri4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	81.2	377	C82876	oligopeptide trans
2	25	78.1	127	G84653	hypothetical prote
3	25	78.1	482	A83855	lysine decarboxyla
4	25	78.1	597	E71130	probable oligopept
5	25	78.1	2796	JC4743	fatty-acid synthas
6	24	75.0	135	T03490	conserved hypothet
7	24	75.0	142	AG2894	hypothetical prote
8	24	75.0	143	G70761	hypothetical prote
9	24	75.0	152	T07858	glycine-rich prote
10	24	75.0	154	C98829	hypothetical prote
11	24	75.0	240	A97826	tRNA/rRNA methyltr
12	24	75.0	287	C81066	conserved hypothet
13	24	75.0	295	JC7926	3-hydroxyisobutyra
14	24	75.0	300	A81804	conserved hypothet
15	24	75.0	302	G86920	probable integral
16	24	75.0	302	E70888	hypothetical prote
17	24	75.0	314	C75003	na+/ca+ exchanging
18	24	75.0	325	D71159	probable Na+/Ca2+-
19	24	75.0	337	T06720	hypothetical prote
20	24	75.0	409	T35155	hypothetical prote
21	24	75.0	453	C83008	conserved hypothet
22	24	75.0	801	T07617	proteinase TMP - t
23	24	75.0	881	T25786	hypothetical prote
24	24	75.0	1016	T05056	hypothetical prote
25	24	75.0	4367	B54802	dynamin heavy chain
26	24	75.0	5032	A35041	ryanodine receptor
27	24	75.0	5035	T146646	ryanodine receptor
28	24	75.0	5037	B35041	ryanodine receptor
29	23	71.9	53	H84193	hypothetical prote

30	23	71.9	55	2	D87423	cytochrome c oxida
31	23	71.9	118	2	T44470	transposase tnnp I
32	23	71.9	133	2	S57885	T-cell receptor al
33	23	71.9	152	2	E70467	hypothetical prote
34	23	71.9	166	2	A69820	hypothetical prote
35	23	71.9	187	2	AD3627	maltose/maltodextr
36	23	71.9	193	2	A82027	probable MAPK alte
37	23	71.9	201	2	T44489	hypothetical prote
38	23	71.9	210	2	D87467	acyl-CoA thioester
39	23	71.9	215	2	T44861	probable hydroxyla
40	23	71.9	229	2	C43431	phosphorylase kina
41	23	71.9	243	2	S43887	restriction endonu
42	23	71.9	243	2	F81130	type II restrictio
43	23	71.9	248	2	D71672	hypothetical prote
44	23	71.9	253	2	D87679	hypothetical prote
45	23	71.9	256	2	A32296	ubiquinol-cytochro

## ALIGNMENTS

### RESULT 1

C82876  
oligopeptide transport system permease protein UUS62 [imported] - Ureaplasma urealyticu  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: C82876  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mi  
A:Reference number: A82870  
A:Accession: C82876  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-377 <GLA>  
A:Cross-references: GB:AF002154; GB:AF222894; NID:G6899557; PIDN:AAF30975.1; GSPDB:GN00  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: oppB; UUS62  
A:Genetic code: SGC3

Query Match 81.2%; Score 26; DB 2; Length 377;  
Best Local Similarity 62.5%; Pred. No. 14;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|||  
DB 146 GVSTSI RP 153

### RESULT 2

G84653  
hypothetical protein At2g25870 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84653  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: G84653  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-127 <STC>  
A:Cross-references: GB:AE002093; NID:G3643596; PIDN:AAC42243.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g25870  
A:Map position: 2

Query Match 78.1%; Score 25; DB 2; Length 127;  
Best Local Similarity 62.5%; Pred. No. 8.8;



Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 19 GVSSVIRP 26

Db

RESULT 3  
 A:Residues: 1-482 <STO>  
 A:Cross-references: GB:AP001507; GB:BA000004; NID:gl0172612; PIDN:BA03760.1; GSPDS:GN00  
 A:Experimental source: strain C-125  
 A:Genetics:  
 A:Gene: BH0041  
 A:Superfamily: lysine decarboxylase

Query Match 78.1%; Score 25; DB 2; Length 482;  
 Best Local Similarity 62.5%; Pred. No. 35;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 173 GVARSIRP 180

Db

RESULT 4  
 B71130  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 A:Accession: B71130  
 R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; PMID:9679194  
 A:Accession: B71130  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-597 <KAW>  
 A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29900.1; PID:g3257217  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 A:Genetics:  
 A:Gene: PH0807  
 C:Superfamily: dipeptide transport protein

Query Match 78.1%; Score 25; DB 2; Length 597;  
 Best Local Similarity 62.5%; Pred. No. 44;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 274 GVSVSIRP 281

Db

RESULT 5  
 JC4743  
 fatty acid synthase (EC 2.3.1.85) - Mycobacterium bovis  
 C:Species: Mycobacterium bovis  
 C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 05-May-2000  
 C:Accession: JC4743

R.Fernandes, N.D.; Kolattukudy, P.E.  
 Gene 170, 35-99, 1996  
 A:Title: Cloning, sequencing and characterization of a fatty acid synthase-encoding gen  
 A:Reference number: JC4743; PMID:96200863; PMID:8621098  
 A:Accession: JC4743  
 A:Molecule type: DNA  
 A:Residues: 1-2796 <FER>  
 A:Cross-references: GB:U36763; NID:gl036834; PIDN:AAB03809.1; PID:gl036835  
 A:Note: the source is designated as Mycobacterium tuberculosis var. bovis BCG  
 C:Comment: This enzyme catalyzes both de novo synthesis and chain elongation of fatty  
 C:Genetics:  
 A:Gene: fas  
 A:Superfamily: Mycobacterium tuberculosis fatty-acid synthase  
 C:Keywords: acyltransferase; coenzyme A; phosphopantetheine; phosphoprotein  
 F:2188-2193/Region: nucleotide binding #status predicted  
 F:57/Active site: Ser #status predicted  
 F:1693/Binding site: phosphopantetheine (Ser) (covalent) #status predicted  
 F:2598/Active site: Cys #status predicted

Query Match 78.1%; Score 25; DB 2; Length 2796;  
 Best Local Similarity 62.5%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 2704 GVHTSIRP 2711

Db

RESULT 6  
 T03490  
 conserved hypothetical protein - Rhodobacter capsulatus  
 C:Species: Rhodobacter capsulatus  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 08-Oct-1999  
 A:Accession: T03490  
 R:Vleek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.  
 Proc Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997  
 A:Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB100  
 A:Reference number: Z14955; PMID:97404404; PMID:9256491  
 A:Accession: T03490  
 A:Status: preliminary; translated from GB/EMBL/DDJ  
 A:Molecule type: DNA  
 A:Residues: 1-135 <VLC>  
 A:Cross-references: EMBL:AF010496; NID:g3138256; PIDN:AAC16143.1; PID:g3128291  
 C:Genetics:  
 A:Map position: 1

Query Match 75.0%; Score 24; DB 2; Length 135;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
 61 GVENAIRP 68

Db

RESULT 7  
 AG2994  
 hypothetical protein Atu3561 [imported] - Agrobacterium tumefaciens (strain C58, Dupont  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 A:Accession: AG2994  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClel  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:21608550; PMID:11743193  
 A:Accession: AG2994  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-142 <KUR>

A;Cross-references: GB:AB008689; PIDN:AA144373.1; PID:g17741969; GSPDB:GN00187  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: Atu3561  
 A;Map position: linear chromosome

Query Match 75.0%; Score 24; DB 2; Length 142;  
 Best Local Similarity 62.5%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 |||  
 DB 3 GVTGIRP 10

RESULT 8  
 G70761  
 hypothetical protein Rv1546 - Mycobacterium tuberculosis (strain H37RV)  
 C;Species: Mycobacterium tuberculosis  
 C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C;Accession: G70761  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: G70761  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-143 <COL>  
 A;Cross-references: GB:274020; GB:AL123456; NID:93261584; PIDN:CAA98316.1; PID:g1403496  
 A;Experimental source: strain H37RV  
 C;Genetics:  
 A;Gene: Rv1546

Query Match 75.0%; Score 24; DB 2; Length 143;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 |||  
 DB 90 GVTGIRP 97

RESULT 9  
 T07858  
 glycine-rich protein - tomato  
 C;Species: Lycopersicon esculentum (tomato)  
 C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 08-Oct-1999  
 C;Accession: T07858  
 R;Smith, A.G.  
 submitted to the EMBL Data Library, October 1995  
 A;Description: A glycine-rich gene expressed in the tapetum of tomato.  
 A;Reference number: Z16160  
 A;Accession: T07858  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-152 <SM>  
 A;Cross-references: EMBL:L47626; NID:g1006564; PIDN:AA78253.1; PID:g1006565  
 A;Experimental source: strain cv VF36; clone T0M127  
 C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 75.0%; Score 24; DB 2; Length 152;  
 Best Local Similarity 62.5%; Pred. No. 21;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 |||  
 DB 56 GVRGAIRP 63

RESULT 10  
 C98289  
 hypothetical protein AGR\_L\_2537 [imported] - Agrobacterium tumefaciens (strain C58, Cer  
 C;Species: Agrobacterium tumefaciens  
 C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C;Accession: C98289  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: C98289  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-154 <KUR>  
 A;Cross-references: GB:AB007870; PIDN:AAK9837.1; PID:g15159771; GSPDB:GN00170  
 C;Genetics:  
 A;Gene: AGR\_L\_2537  
 A;Map position: linear chromosome

Query Match 75.0%; Score 24; DB 2; Length 154;  
 Best Local Similarity 62.5%; Pred. No. 21;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 |||  
 DB 15 GVTGIRP 22

RESULT 11  
 A97826  
 tRNA/rRNA methyltransferase (EC 2.1.1.1-) [imported] - Rickettsia conorii (strain Malish  
 C;Species: Rickettsia conorii  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 22-Oct-2001  
 C;Accession: A97826  
 R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R  
 Science 293, 2093-2098, 2001  
 A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.  
 A;Reference number: A97700; MUID:21442074; PMID:11557893  
 A;Accession: A97826  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-240 <KUR>  
 A;Cross-references: GB:AE006914; PIDN:AA03547.1; PID:g15620124; GSPDB:GN00173  
 C;Genetics:  
 A;Gene: RCI009  
 A;Superfamily: conserved hypothetical protein HI0860  
 C;Keywords: methyltransferase

Query Match 75.0%; Score 24; DB 2; Length 240;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 |||  
 DB 72 GIAXIRP 79

RESULT 12  
 C81066  
 conserved hypothetical protein NMB1593 [imported] - Neisseria meningitidis (strain MC58  
 C;Species: Neisseria meningitidis  
 C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C;Accession: C81066  
 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.  
 Hickey, E.K.; Hatt, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.  
 et al.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.  
 Science 287, 1809-1815, 2000  
 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.  
 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
 A;Reference number: A81000; MUID:20175755; PMID:10710307  
 A;Accession: C81066  
 A;Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-287 <RET>  
A:Cross-references: GB:AE002509; GB:AE002098; NID:g7226832; PIDN:AAF41946.1; PID:g722684  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1593

Query Match 75.0%; Score 24; DB 2; Length 287;  
Best Local Similarity 50.0%; Pred. No. 41;  
Matches 4; Conservative 1; Mismatches 0; Gaps 0;  
Indels 3;

QY 3 GVXXXIRP 10  
|: |||  
Db 266 GILSGIRP 273

RESULT 13  
JC7926  
3-hydroxyisobutyrate dehydrogenase (EC 1.1.1.31) - Pseudomonas putida E23  
N:Alternate names: 3-hydroxy-2-methylpropionate: NAD+ oxidoreductase; NAD+ -dependent 3-  
C:Species: Pseudomonas putida E23  
C>Date: 31-Mar-2003 #sequence\_revision 31-Mar-2003 #text\_change 14-Jul-2003  
C:Accession: JC7926; PC7219  
R:Chowdhury, E.K.; Akaiishi, Y.; Nagata, S.; Misono, H.  
Bioosci. Biotechnol. Biochem. 67, 438-441, 2003  
A:Title: Cloning and overexpression of the 3-hydroxyisobutyrate dehydrogenase gene from  
A:Reference number: JC7926  
A:Accession: JC7926  
A:Molecule type: DNA  
A:Residues: 1-295 <CHO>  
A:Cross-references: DDBJ:AB090855  
A:Accession: PC7219  
A:Molecule type: protein  
A:Residues: 1-33;34-51;246-265;293-293 <CH2>  
C:Comment: This enzyme, which is in the 3-hydroxyacid dehydrogenase family, catalyzes the  
alane metabolism and is specific for the L-enantiomer of 3-hydroxyisobutyrate.  
C:Keywords: oxidoreductase

Query Match 75.0%; Score 24; DB 2; Length 295;  
Best Local Similarity 62.5%; Pred. No. 42;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|: |||  
Db 79 GVLGIRP 86

RESULT 14  
A81804  
conserved hypothetical integral membrane protein NMA1785 [imported] - Neisseria meningitidis  
C:Species: Neisseria meningitidis  
C>Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: A81804  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A:Reference number: A81775; MUID:20222556; PMID:10761919  
A:Accession: A81804  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-300 <PAR>  
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CA885013.1; PID:g738042  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA1785

Query Match 75.0%; Score 24; DB 2; Length 300;  
Best Local Similarity 50.0%; Pred. No. 43;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|: |||

Db 279 GILSSIRP 286

RESULT 15  
G86920  
probable integral membrane protein [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: G86920  
R:Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: G86920  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-302 <STO>  
A:Cross-references: GB:AL450380; NID:gl3092477; PIDN:CA29603.1; GSPDB:GN00147  
C:Genetics:  
A:Gene: ML0095

Query Match 75.0%; Score 24; DB 2; Length 302;  
Best Local Similarity 62.5%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10  
|: |||  
Db 16 GVKAIRP 23

Search completed: April 1, 2004, 17:40:43  
Job time: 12.9649 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:29:18 ; Search time 6.75439 Seconds  
(without alignments)  
84.800 Million cell updates/sec

Title: US-09-833-196-6  
Perfect score: 32  
Sequence: 1 XXGVXXXIRPX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	75.0	143	1 YF46 MYCTU	Q10780 mycobacteri
2	24	75.0	198	1 YF2 ARATH	Q04653 arabidopsis
3	24	75.0	4349	1 DYHC FUSSO	P78716 fusarium so
4	24	75.0	4367	1 DYHC NEUCR	P45443 neurospora
5	24	75.0	5035	1 RYRI PIG	P16960 sus scrofa
6	24	75.0	5037	1 RYRI RABIT	P11716 oryctolagus
7	24	75.0	5038	1 RYRI HUMAN	P21817 homo sapien
8	23	71.9	126	1 YJ53 AQUAE	O67766 aquifex aeo
9	23	71.9	200	1 ORN STRAW	Q82688 streptomyce
10	23	71.9	200	1 ORN STROCO	P57666 streptomyce
11	23	71.9	201	1 ORN STRGR	P57667 streptomyce
12	23	71.9	213	1 VATD THEAC	Q9hm63 thermoplas
13	23	71.9	243	1 T2N4 NEILA	P50183 neisseria l
14	23	71.9	256	1 UCRI RAT	P20788 rattus norv
15	23	71.9	274	1 UCRI BOVIN	P13272 bos taurus
16	23	71.9	274	1 UCRI HUMAN	P47985 homo sapien
17	23	71.9	320	1 MDCF RHIME	P56943 rhizobium m
18	23	71.9	321	1 AX18 ARATH	Q38826 arabidopsis
19	23	71.9	354	1 LEU3 BACAA	Q81t67 bacillus an
20	23	71.9	354	1 LEU3 BHACH	Q81g11 bacillus ce
21	23	71.9	380	1 LSU3 PACHU	O59930 phanerocha
22	23	71.9	401	1 SP2P PACSU	P37968 bacillus su
23	23	71.9	424	1 CLPX RAUSO	Q8xy96 ralstonia s
24	23	71.9	495	1 NUSA ECOLI	P03003 escherichia
25	23	71.9	500	1 NUSA SALT	P37430 salmonella
26	23	71.9	850	1 RN12 YEAST	P32843 saccharomyc
27	23	71.9	885	1 RPAL THEAC	Q03585 thermoplas
28	23	71.9	1115	1 PAN2 YEAST	P53010 saccharomyc
29	23	71.9	1223	1 KPB1 HUMAN	P46020 homo sapien
30	23	71.9	1235	1 KPB2 HUMAN	P46019 homo sapien
31	23	71.9	1235	1 KPB2 MOUSE	Q8bwj3 mus musculu
32	23	71.9	1237	1 KPB1 RABIT	P18688 oryctolagus
33	23	71.9	1241	1 KPB1_MOUSE	P18826 mus musculu

34	23	71.9	1242	1 KPB1 RAT	O64649 rattus norv
35	23	71.9	1538	1 PP8E MYCTU	Q10978 mycobacteri
36	23	71.9	1541	1 NK3A HUMAN	Q9y4c0 homo sapien
37	23	71.9	1620	1 PCX3 MOUSE	Q8vi59 mus musculu
38	22	68.8	49	1 COXH THOJB	P80977 thunnus obe
39	22	68.8	135	1 NUSB WOLSU	Q7mae9 wolinnella s
40	22	68.8	141	1 UL41 HCMVA	P16814 human cytom
41	22	68.8	160	1 UL60 HCMVA	P16817 human cytom
42	22	68.8	170	1 C13 MAIZE	P33050 zea mays (m
43	22	68.8	175	1 HES3 MOUSE	O61657 mus musculu
44	22	68.8	219	1 ISPD BACTIN	Q8a0u8 bacteroides
45	22	68.8	237	1 PYRF LACPL	P77888 lactobacill

## ALIGNMENTS

### RESULT 1

ID	YF46 MYCTU	STANDARD	PRT	143 AA
AC	Q10780;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Hypothetical protein Rv1546/MT1597/MB1573.			
GN	Rv1546 OR MT1597 OR MTCY48.19C OR MB1573.			
OS	Mycobacterium tuberculosis, and			
OS	Mycobacterium bovis.			
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;			
OC	Corynebacteriineae; Mycobacteriaceae; Mycobacterium.			
OX	NCBI_TaxID=1773, 1765;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.tuberculosis; STRAIN=H37Rv;			
RX	MEDLINE=98295987; PubMed=9634230;			
RA	Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,			
RA	Gordon S.V., Eiglmeier K., Brown D., Gas S., Barry C.E. III, Tekala F.,			
RA	Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,			
RA	Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,			
RA	Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,			
RA	Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,			
RA	Rutter S., Seeger K., Skelton S., Squares S., Squares R.,			
RA	Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;			
RT	"Deciphering the biology of Mycobacterium tuberculosis from the			
RT	complete genome sequence."			
RL	Nature 393:537-544(1998).			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;			
RX	MEDLINE=22206494; PubMed=12218036;			
RA	Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,			
RA	Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,			
RA	Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,			
RA	Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,			
RA	Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;			
RT	"Whole-genome comparison of Mycobacterium tuberculosis clinical and			
RT	laboratory strains."			
RL	J. Bacteriol. 184:5479-5490(2002).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=M.bovis; STRAIN=AF122/97;			
RX	MEDLINE=22709107; PubMed=12789972;			
RA	Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,			
RA	Pryor M., Duthoy R., Grondin S., Lacroix C., Monsemp C., Simon S.,			
RA	Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,			
RA	Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;			
RT	"The complete genome sequence of Mycobacterium bovis."			
RL	Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Z74020; CAA98316.1; --  
 DR EMBL; AE007026; AAK45864.1; --  
 DR EMBL; BX248339; CAD96240.1; --  
 DR PIR; G70761; G70761.  
 DR TIGR; MT1597; --

KW Tuberculin; RV1546; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 143 AA; 15324 MW; D3E4CAA547C48BB3 CRC64;

Query Match 75.0%; Score 24; DB 1; Length 143;  
 Best Local Similarity 50.0%; Pred. No. 14;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 3 GVXXXIRP 10  
 |||||  
 Db 90 GVITVAP 97

RESULT 2  
 ID IF22 ARATH STANDARD; PRT; 198 AA.  
 AC O04653; O9SANA; --  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Eukaryotic translation initiation factor 4E-2 (EIF4E-2) (EIF4E-2)  
 DE (mRNA cap-binding protein) (EIF-(iso)4F 25 kDa subunit) (EIF-(iso)4F  
 DE p28 subunit) (EIF4Eiso protein) (EIF(iso)4E).  
 GN EIF4E2 OR AT5G35620 OR MJB4.8.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Arabidopsis thaliana; Embryophyta; Tracheophyta;  
 CC Eukaryota; Viridiplantae; Streptophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
 ON NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.  
 RX MEDLINE=97379268; PubMed=9234949;  
 RA Wittmann S., Chatel H., Fortin M.G., Laliberte J.F.;  
 RT "Interaction of the viral protein genome linked of turnip mosaic  
 RT potyvirus with the translational eukaryotic initiation factor (iso)  
 RT 4E of Arabidopsis thaliana using the yeast two-hybrid system.";  
 RL Virology 234:84-92(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Rodriguez C.M., Preire M.A., Robaglia C.;  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=98403884; PubMed=9734815;  
 RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,  
 RA Tabata S.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.  
 RT Sequence features of the regions of 1,367,185 bp covered by 19  
 RT physically assigned P1 and TAC clones.";  
 RL DNA Res. 5:203-216(1998).

CC -!- FUNCTION: Recognizes and binds the 7-methylguanosine-containing  
 CC mRNA cap during an early step in the initiation of protein  
 CC synthesis and facilitates ribosome binding by inducing the  
 CC unwinding of the mRNAs secondary structures.  
 CC -!- SUBUNIT: EIF4F is a multi-subunit complex, the composition of  
 CC which varies with external and internal environmental conditions.  
 CC It is composed of at least EIF4A, EIF4E and EIF4G. EIF4E is also  
 CC known to interact with other partners. In higher plants two  
 CC isoforms of EIF4F have been identified, named EIF4F and  
 CC EIF(iso)4F. EIF4F has subunits p220 and p28, whereas EIF(iso)4F  
 CC has subunits p82 and p26. This isoform interacts with the viral  
 CC protein genome linked (VPg)-proteinase of turnip mosaic potyvirus.

-!- TISSUE SPECIFICITY: Abundant in floral organs and in young

CC developing tissues.  
 CC -!- SIMILARITY: Belongs to the eukaryotic initiation factor 4E family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U62044; AAB66906.1; ALT\_INIT.  
 DR EMBL; Y10547; CAA71579.1; --  
 DR EMBL; AB013393; BAB09303.1; --  
 DR HSSP; P07260; 1AP8  
 DR InterPro; IPR001040; TIF\_eIF\_4E.  
 DR Pfam; PF01652; IF4E; 1.  
 DR ProDom; PD003697; TIF\_eIF\_4E; 1.  
 DR PROSITE; PS00813; IF4E; 1.  
 KW Protein biosynthesis; Translation regulation; Initiation factor;  
 KW RNA-binding; Multigene family.  
 SQ SEQUENCE 198 AA; 22514 MW; 71FEB309E073A9D2 CRC64;

Query Match 75.0%; Score 24; DB 1; Length 198;  
 Best Local Similarity 50.0%; Pred. No. 19;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 3 GVXXXIRP 10  
 |||||  
 Db 139 GVASVRP 146

RESULT 3  
 ID DYHC FUSO STANDARD; PRT; 4349 AA.  
 AC P78716;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Dynein heavy chain, cytosolic (DYHC).  
 GN DHCL.  
 OS Fusarium solani (subsp. pisi) (Nectria haematococca).  
 OS Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.  
 ON NCBI\_TaxID=70791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=T213;  
 RA Inoue S., Aist J.R., Turgeon B.G., Yoder O.C.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular  
 CC retrograde motility of vesicles and organelles along microtubules.  
 CC Dynein has ATPase activity; the force-producing power stroke is  
 CC thought to occur on release of ADP.  
 CC -!- SUBUNIT: Consists of at least two heavy chains and a number of  
 CC intermediate and light chains.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem  
 CC (which binds cargo and interacts with other dynein components),  
 CC and the head or motor domain. The motor contains six tandemly-  
 CC linked AAA domains in the head, which form a ring. A stalk-like  
 CC structure (formed by two of the coiled coil domains) protrudes  
 CC between AAA 4 and AAA 5 and terminates in a microtubule-binding  
 CC site. A seventh domain may also contribute to this ring; it is not  
 CC clear whether the N-terminus or the C-terminus forms this extra  
 CC domain. There are four well-conserved and two non-conserved ATPase  
 CC sites, one per AAA domain. Probably only one of these (within AAA  
 CC 1) actually hydrolyzes ATP, the others may serve a regulatory  
 CC function.

CC -!- SIMILARITY: Belongs to the dynein heavy chain family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

intermediate and light chains.  
-1- SUBCELLULAR LOCATION: Cytoplasmic.  
-1- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem (which binds cargo and interacts with other dynein components), and the head or motor domain. The motor contains six tandemly-linked AAA domains in the head, which form a ring. A stalk-like structure (formed by two of the coiled coil domains) protrudes between AAA 4 and AAA 5 and terminates in a microtubule-binding site. A seventh domain may also contribute to this ring; it is not clear whether the N-terminus or the C-terminus forms this extra domain. There are four well-conserved and two non-conserved ATPase sites, one per AAA domain. Probably only one of these (within AAA 1) actually hydrolyzes ATP, the others may serve a regulatory function.

-1- SIMILARITY: Belongs to the dynein heavy chain family.  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L31504; AAC33176.1; -  
PIR: B54802; B54802.  
InterPro: IPR003593; AAA\_ATPase.  
InterPro: IPR004273; Dynein\_heavy.  
Pfam: PF03028; Dynein\_heavy; 1.  
SMART; SM00382; AAA; 3.  
Motor protein; Dynein; Microtubule; ATP-binding; Repeat; Coiled coil.  
DOMAIN 1 1904 STEM (BY SIMILARITY).  
DOMAIN 1905 2130 AAA 1 (BY SIMILARITY).  
DOMAIN 2202 2460 AAA 2 (BY SIMILARITY).  
DOMAIN 2566 2815 AAA 3 (BY SIMILARITY).  
DOMAIN 2909 3179 AAA 4 (BY SIMILARITY).  
DOMAIN 3193 3481 STALK (BY SIMILARITY).  
DOMAIN 3565 3794 AAA 5 (BY SIMILARITY).  
DOMAIN 4003 4215 AAA 6 (BY SIMILARITY).  
DOMAIN 676 693 COILED COIL (POTENTIAL).  
DOMAIN 1176 1215 COILED COIL (POTENTIAL).  
DOMAIN 1327 1351 COILED COIL (POTENTIAL).  
DOMAIN 1557 1574 COILED COIL (POTENTIAL).  
DOMAIN 1637 1668 COILED COIL (POTENTIAL).  
DOMAIN 2195 2218 COILED COIL (POTENTIAL).  
DOMAIN 3193 3296 COILED COIL (POTENTIAL).  
DOMAIN 3423 3481 COILED COIL (POTENTIAL).  
DOMAIN 3778 3809 COILED COIL (POTENTIAL).  
NP\_BIND 1943 1950 ATP (POTENTIAL).  
NP\_BIND 2240 2247 ATP (POTENTIAL).  
NP\_BIND 2605 2612 ATP (POTENTIAL).  
NP\_BIND 2947 2954 ATP (POTENTIAL).  
SEQUENCE 4367 AA; 495568 MW; 1E10F3E2D170D6DF CRC64;

Query Match 75.0%; Score 24; DB 1; Length 4367;  
Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 GVXXXIRP 10  
DB 2815 GVYEAIRP 2822

RESULT 5  
RYR1\_PIG  
ID\_RYR1\_PIG STANDARD; PRT; 5035 AA.  
AC P16960;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RyR1) (RyR1-1) (Skeletal muscle calcium release channel).  
GN RYR1 OR CRC.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U84215; AAC33176.1; -  
InterPro: IPR003593; AAA\_ATPase.  
InterPro: IPR004273; Dynein\_heavy.  
Pfam; PF03028; Dynein\_heavy; 1.  
SMART; SM00382; AAA; 3.  
Motor protein; Dynein; Microtubule; ATP-binding; Repeat; Coiled coil.  
DOMAIN 1 1907 STEM (BY SIMILARITY).  
DOMAIN 1908 2133 AAA 1 (BY SIMILARITY).  
DOMAIN 2201 2459 AAA 2 (BY SIMILARITY).  
DOMAIN 2565 2814 AAA 3 (BY SIMILARITY).  
DOMAIN 2908 3177 AAA 4 (BY SIMILARITY).  
DOMAIN 3186 3477 STALK (BY SIMILARITY).  
DOMAIN 3563 3792 AAA 5 (BY SIMILARITY).  
DOMAIN 4001 4213 AAA 6 (BY SIMILARITY).  
DOMAIN 459 480 COILED COIL (POTENTIAL).  
DOMAIN 1178 1215 COILED COIL (POTENTIAL).  
DOMAIN 1266 1293 COILED COIL (POTENTIAL).  
DOMAIN 1334 1354 COILED COIL (POTENTIAL).  
DOMAIN 1560 1577 COILED COIL (POTENTIAL).  
DOMAIN 1640 1670 COILED COIL (POTENTIAL).  
DOMAIN 2194 2217 COILED COIL (POTENTIAL).  
DOMAIN 3186 3294 COILED COIL (POTENTIAL).  
DOMAIN 3420 3477 COILED COIL (POTENTIAL).  
DOMAIN 3774 3807 COILED COIL (POTENTIAL).  
NP\_BIND 1946 1953 ATP (POTENTIAL).  
NP\_BIND 2239 2246 ATP (POTENTIAL).  
NP\_BIND 2604 2611 ATP (POTENTIAL).  
NP\_BIND 2946 2953 ATP (POTENTIAL).  
SEQUENCE 4349 AA; 493453 MW; FCB3C7152B36A1BF CRC64;

Query Match 75.0%; Score 24; DB 1; Length 4349;  
Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 GVXXXIRP 10  
DB 2814 GVYEAIRP 2821

RESULT 4  
DYHC\_NEUCR  
ID\_DYHC\_NEUCR STANDARD; PRT; 4367 AA.  
AC P45443;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Dynein heavy chain, cytosolic (DYHC).  
GN RO-1.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A / FGSC 987;  
RX MEDLINE=95014704; PubMed=7929559;  
RA Plamann M., Minke P.F., Tinsley J.H., Bruno K.S.;  
RT "Cytoplasmic dynein and actin-related protein Arp1 are required for normal nuclear distribution in filamentous fungi";  
RL J. Cell Biol. 127:139-149(1994).  
CC -1- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular retrograde motility of vesicles and organelles along microtubules. Dynein has ATPase activity; the force-producing power stroke is thought to occur on release of ADP. Required to maintain uniform nuclear distribution in hyphae.  
CC -1- SUBUNIT: Consists of at least two heavy chains and a number of

Query Match 75.0%; Score 24; DB 1; Length 4367;  
Best Local Similarity 62.5%; Pred. No. 3.6e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 3 GVXXXIRP 10  
DB 2815 GVYEAIRP 2822

RESULT 5  
RYR1\_PIG  
ID\_RYR1\_PIG STANDARD; PRT; 5035 AA.  
AC P16960;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RyR1) (RyR1-1) (Skeletal muscle calcium release channel).  
GN RYR1 OR CRC.

OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Norwegian landrace; TISSUE=Skeletal muscle;  
 RA MEDLINE=93036581; PubMed=1329581;  
 RX Harbitz I., Kristensen T., Bones M., Kran S., Davies M.;  
 RT "DNA sequence of the skeletal muscle calcium release channel cDNA and  
 RT verification of the Arg615-->Cys615 mutation, associated with porcine  
 RT malignant hyperthermia, in Norwegian landrace pigs.";  
 RL Anim. Genet. 23:395-402(1992).  
 RN [2]  
 RP SEQUENCE OF 1129-2801 FROM N.A.  
 RA Brenig B.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1129-2643 FROM N.A.  
 RC STRAIN-German Landrace; TISSUE=Liver;  
 RX MEDLINE=94117003; PubMed=828238;  
 RA Leeb T., Schmolzi S., Brem G., Brenig B.;  
 RT "Genomic organization of the porcine skeletal muscle ryanodine  
 RT receptor (RYR1) gene coding region 4624 to 7929.";  
 RL Genomics 18:349-354(1993).  
 RN [4]  
 RP SEQUENCE OF 4785-5035 FROM N.A.  
 RX MEDLINE=91065640; PubMed=2174405;  
 RA Harbitz I., Chowdhary B., Thomsen P.D., Davies W., Kaufman U.,  
 RA Kran S., Gustavsson I., Christensen K., Hauge J.G.;  
 RT "Assignment of the porcine calcium release channel gene, a candidate  
 RT for the malignant hyperthermia locus, to the 6p11--q21 segment of  
 RT chromosome 6.";  
 RL Genomics 8:243-248(1990).  
 CC -!- FUNCTION: Communication between transverse-tubules and  
 CC sarcoplasmic reticulum. Contraction of skeletal muscle is  
 CC triggered by release of calcium ions from SR following  
 CC depolarization of T-tubules.  
 CC -!- SUBUNIT: Homotetramer (potential).  
 CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium  
 CC ions, magnesium ions, ATP and calmodulin.  
 CC -!- MISCELLANEOUS: The calcium release channel activity resides in the  
 CC C-terminal region while the remaining part of the protein  
 CC constitutes the 'foot' structure spanning the junctional gap  
 CC between the SR and the T-tubule. It is possible that the foot  
 CC structure interacts with the cytoplasmic region of the  
 CC dihydropyridine receptor.  
 CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-  
 CC release channel in junctional SR and modulates its activity.  
 CC -!- SIMILARITY: Belongs to the ryanodine receptor family.  
 CC -!- SIMILARITY: Contains 5 MIR domains.  
 CC -!- SIMILARITY: Contains 3 SPRY domains.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; X52880; CAA44674.1; ALT\_SEQ.  
 DR EMBL; X68247; CAA48318.1; -  
 DR EMBL; X69465; CAA49225.1; -  
 DR EMBL; W32501; AAA31022.1; -  
 DR InterPro; IPR000659; Ca-rel channel.  
 DR InterPro; IPR001682; Ca/Na pore.  
 DR InterPro; IPR002048; EF-hand.  
 DR InterPro; IPR005821; Ion\_trans.  
 DR InterPro; IPR003608; MIR.  
 DR InterPro; IPR001215; Ryanodn\_receptor.  
 DR InterPro; IPR003032; RYR.  
 DR InterPro; IPR003877; SPRY\_receptor.

DR Pfam; PF005020; ion\_trans; 1.  
 DR Pfam; PF02815; MIR\_4.  
 DR Pfam; PF01365; RYR\_ITPR; 2.  
 DR Pfam; PF02026; RYR; 4.  
 DR Pfam; PF00622; SPRY; 3.  
 DR PRINTS; PR00795; RYANODINER.  
 DR SMART; SM00472; MIR; 4.  
 DR SMART; SM00449; SPRY; 3.  
 DR PROSITE; PS00919; MIR; 5.  
 KW Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;  
 KW Phosphorylation; Glycoprotein; CYTOPLASMIC.  
 FT DOMAIN 1 3124  
 FT TRANSMEM 3125 3145 M' (POTENTIAL).  
 FT TRANSMEM 3189 3207 M'' (POTENTIAL).  
 FT TRANSMEM 3980 3999 M1 (POTENTIAL).  
 FT TRANSMEM 4018 4036 M2 (POTENTIAL).  
 FT TRANSMEM 4274 4297 M3 (POTENTIAL).  
 FT TRANSMEM 4339 4359 M4 (POTENTIAL).  
 FT TRANSMEM 4557 4578 M5 (POTENTIAL).  
 FT TRANSMEM 4646 4669 M6 (POTENTIAL).  
 FT TRANSMEM 4787 4807 M7 (POTENTIAL).  
 FT TRANSMEM 4835 4854 M8 (POTENTIAL).  
 FT TRANSMEM 4877 4896 M9 (POTENTIAL).  
 FT TRANSMEM 4921 4935 M10 (POTENTIAL).  
 FT DOMAIN 98 153 MIR 1.  
 FT DOMAIN 160 205 MIR 2.  
 FT DOMAIN 211 265 MIR 3.  
 FT DOMAIN 271 334 MIR 4.  
 FT DOMAIN 336 393 MIR 5.  
 FT DOMAIN 559 797 SPRY 1.  
 FT DOMAIN 1085 1208 SPRY 2.  
 FT DOMAIN 1430 1570 SPRY 3.  
 FT DOMAIN 1874 1925 GLU-RICH (ACIDIC).  
 FT DOMAIN 4458 4526 PRO-RICH.  
 FT DOMAIN 842 2960 6 X APPROXIMATE REPEATS.  
 FT REPEAT 842 955 1.  
 FT REPEAT 956 1069 2.  
 FT REPEAT 1345 1360 3 (INCOMPLETE).  
 FT REPEAT 1373 1388 4 (INCOMPLETE).  
 FT REPEAT 2727 2846 5.  
 FT REPEAT 2847 2960 6.  
 FT MOD\_RES 2844 2844  
 FT MOD\_RES 3947 3947  
 FT MOD\_RES 4320 4320  
 FT CARBOHYD 3467 3467 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 3475 3475 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 3904 3904 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 3945 3945 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 4144 4144 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 4862 4862 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 2092 2092 A -> P (IN REF. 2 AND 3).  
 SQ SEQUENCE 5035 AA; 565317 MW; E00613F2027B94A4 CRC64;  
 Query Match 75.0%; Score 24; DB 1; Length 5035;  
 Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXXIRP 10  
 DB 1767 GVITSLRP 1774  
 RESULT 6  
 ID RYR1 RABIT STANDARD; PRT; 5037 AA.  
 AC P11716;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ryanodine receptor 1 (skeletal muscle-type ryanodine receptor) (RYR1)  
 DE (RYR-1) (skeletal muscle calcium release channel).  
 GN RYR1.

DR Pfam; PF02815; MIR; 4.  
DR Pfam; PF01365; RYDR\_ITPR; 2.  
DR Pfam; PF02026; RYR; 4.  
DR Pfam; PF00622; SPRY; 3.  
DR PRINTS; PRO0795; RYANODINER.  
DR SMART; SM00472; MIR; 4.  
DR SMART; SM00449; SPRY; 3.  
DR PROSITE; PS50919; MIR; 5.  
KW Receptor, Transmembrane; Ionic channel; Calcium channel; Repeat;  
KW Phosphorylation; Glycoprotein.  
FT DOMAIN 1 3123 CYTOPLASMIC.  
FT TRANSMEM 3124 3144 M' (POTENTIAL).  
FT TRANSMEM 3188 3206 M'' (POTENTIAL).  
FT TRANSMEM 3985 4004 M1 (POTENTIAL).  
FT TRANSMEM 4023 4041 M2 (POTENTIAL).  
FT TRANSMEM 4277 4300 M3 (POTENTIAL).  
FT TRANSMEM 4342 4362 M4 (POTENTIAL).  
FT TRANSMEM 4559 4580 M5 (POTENTIAL).  
FT TRANSMEM 4648 4671 M6 (POTENTIAL).  
FT TRANSMEM 4789 4809 M7 (POTENTIAL).  
FT TRANSMEM 4837 4856 M8 (POTENTIAL).  
FT TRANSMEM 4879 4898 M9 (POTENTIAL).  
FT TRANSMEM 4914 4937 M10 (POTENTIAL).  
FT DOMAIN 98 153 MIR 1.  
FT DOMAIN 160 205 MIR 2.  
FT DOMAIN 211 265 MIR 3.  
FT DOMAIN 271 334 MIR 4.  
FT DOMAIN 336 394 MIR 5.  
FT DOMAIN 659 797 SPRY 1.  
FT DOMAIN 1085 1208 SPRY 2.  
FT DOMAIN 1430 1570 SPRY 3.  
FT DOMAIN 1873 1913 GLU-RICH (ACIDIC).  
FT DOMAIN 842 2959 6 X APPROXIMATE REPEATS.  
FT REPEAT 842 955 1.  
FT REPEAT 956 1069 2.  
FT REPEAT 1345 1360 3 (INCOMPLETE).  
FT REPEAT 1373 1388 4 (INCOMPLETE).  
FT REPEAT 2726 2845 5.  
FT REPEAT 2846 2959 6.  
FT MOD\_RES 2843 2843 PHOSPHORYLATION (BY PKA AND PKG).  
FT MOD\_RES 3952 3952 PHOSPHORYLATION (POTENTIAL).  
FT MOD\_RES 4323 4323 PHOSPHORYLATION (POTENTIAL).  
FT CARBOHYD 3466 3466 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 3909 3909 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 3950 3950 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 4149 4149 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 4864 4864 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 2015 2015 E -> D (IN REF. 2).  
FT CONFLICT 3481 3485 MISSING (IN REF. 2).  
FT SEQUENCE 5037 AA; 565238 MW; 4ABD87AA26697070 CRC64;  
Query Match 75.0%; Score 24; DB 1; Length 5037;  
Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps  
QY 3 GVXXXIRP 10  
DB 1766 GVTTSLRP 1773  
RESULT 7  
RYR1\_HUMAN  
AC ID RYR1\_HUMAN STANDARD; PRT; 5038 AA.  
AD C21817; Q16314; Q16368; Q9NPK1; Q9P1U4;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RyR1).  
DE (RyR-1) (Skeletal muscle calcium release channel).  
OS Homo sapiens (Human).  
GN RYR1 OR RYDR.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.



OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2), AND PARTIAL SEQUENCE.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=90130482; PubMed=2298749;  
RA Zorzato F., Fujii J., Otsu K., Phillips M.S., Green N.M., Lai F.A.,  
RA Meisner G., McLennan D.H.;  
RT "Molecular cloning of cDNA encoding human and rabbit forms of the  
RT Ca2+ release channel (ryanodine receptor) of skeletal muscle  
RT sarcoplasmic reticulum";  
RL J. Biol. Chem. 265:2244-2256 (1990).  
RN [2]  
RP REVISIONS TO 2324; 2840 AND 3380, VARIANT MH ARG-248, AND VARIANTS  
RP CYS-471; LEU-1787 AND CYS-2060.  
RC TISSUE=Muscle;  
RX MEDLINE=92372020; PubMed=1354642;  
RA Gillard B.F., Otsu K., Fujii J., Duff C.L., de Leon S., Khanna V.K.,  
RA Britt B.A., Worton R.G., McLennan D.H.;  
RA "Polymorphisms and deduced amino acid substitutions in the coding  
RT sequence of the ryanodine receptor (RYR1) gene in individuals with  
RT malignant hyperthermia";  
RL Genomics 13:1247-1254 (1992).  
RN [3]  
RP REVISIONS TO 1365-1368, VARIANT CCD/MH HIS-2435, AND ALTERNATIVE  
RP SPLICING.  
RC TISSUE=Muscle;  
RX MEDLINE=94035117; PubMed=8220422;  
RA Zhang Y., Chen H.S., Khanna V.K., de Leon S., Phillips M.S.,  
RA Schappert K., Britt B.A., Brownell A.K.W., McLennan D.H.;  
RA "A mutation in the human ryanodine receptor gene associated with  
RT central core disease";  
RL Nat. Genet. 5:46-50 (1993).  
RN [4]  
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.  
RX MEDLINE=96299657; PubMed=8661021;  
RA Phillips M.S., Fujii J., Khanna V.K., de Leon S., Yokobata K.,  
RA de Jong P.J., McLennan D.H.;  
RT "The structural organization of the human skeletal muscle ryanodine  
RT receptor (RYR1) gene";  
RL Genomics 34:24-41 (1996).  
RN [5]  
RP SEQUENCE OF 598-722 FROM N.A.  
RC TISSUE=Skeletal muscle;  
RX MEDLINE=92347889; PubMed=1639409;  
RA Otsu K., Phillips M.S., Khanna V.K., de Leon S., McLennan D.H.;  
RA "Refinement of diagnostic assays for a probable causal mutation for  
RT porcine and human malignant hyperthermia";  
RL Genomics 13:835-837 (1992).  
RN [6]  
RP SEQUENCE OF 3339-5038 FROM N.A. (ISOFORM 2).  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Velasco N., Do L., Regala W., Terry A., Ganes J.,  
RA Dangnan L., Eriar A., Christensen M., Georgescu A., Avila J.,  
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,  
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,  
RA Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S.,  
RA Kobayashi A., Olsen A.S., Carrano A.V.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 4596-4974 FROM N.A.  
RC TISSUE=Myotetrium;  
RX MEDLINE=96032536; PubMed=7556644;  
RA Lynn S., Morgan J.M., Lamb H.K., Meisner G., Gillespie J.I.;  
RT "Isolation and partial cloning of ryanodine-sensitive Ca2+ release  
RT channel protein isoforms from human myometrial smooth muscle";  
RL FEBS Lett. 372:6-12 (1995).  
RN [8]  
RP TISSUE SPECIFICITY.  
RX MEDLINE=98268728; PubMed=9607712;  
RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;  
RT "Partial cloning and differential expression of ryanodine  
RT receptor/calcium-release channel genes in human tissues including

the hippocampus and cerebellum";  
RN Neurosci 85:205-216 (1998).  
RN [9]  
RP S-NITROSYLATION OF CYS-3635.  
RX MEDLINE=21457268; PubMed=11562475;  
RA Sun J., Xin C., Eu J.P., Stamler J.S., Meisner G.;  
RA "Cysteine-3635 is responsible for skeletal muscle ryanodine receptor  
RT modulation by NO";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:11158-11162 (2001).  
RN [10]  
RP VARIANT MH CYS-614.  
RX MEDLINE=92128955; PubMed=1774074;  
RA Gillard B.F., Otsu K., Fujii J., Khanna V.K., de Leon S.,  
RA Derdenezi J., Britt B.A., Duff C.L., Worton R.G., McLennan D.H.;  
RA "A substitution of cysteine for arginine 614 in the ryanodine  
RT receptor is potentially causative of human malignant hyperthermia";  
RL Genomics 11:751-755 (1991).  
RN [11]  
RP VARIANTS CCD/MH CYS-163 AND MET-403.  
RX MEDLINE=94035119; PubMed=8220423;  
RA Quane K.A., Healy J.M.S., Keating K.E., Manning B.M., Couch F.J.,  
RA Palmucci L.M., Doriguzzi C., Fagerlund T.H., Berg K., Ordling H.,  
RA Bendixen D., Mortier W., Linz U., Muller C.R., McCarthy T.V.;  
RT "Mutations in the ryanodine receptor gene in central core disease and  
RT malignant hyperthermia";  
RL Nat. Genet. 5:51-55 (1993).  
RN [12]  
RP VARIANT CCD/MH SER-522.  
RX MEDLINE=95130087; PubMed=7829078;  
RA Quane K.A., Keating K.E., Healy J.M.S., Manning B.M.,  
RA Krivosic-Horber R., Krivosic I., Monnier N., Lunardi J.,  
RA McCarthy T.V.;  
RT "Mutation screening of the RYR1 gene in malignant hyperthermia:  
RT detection of a novel Tyr to Ser mutation in a pedigree with  
RT associated central cores";  
RL Genomics 23:236-239 (1994).  
RN [13]  
RP VARIANT MH ARG-341.  
RX MEDLINE=94282042; PubMed=8012359;  
RA Quane K.A., Keating K.E., Manning B.M., Healy J.M.S., Monniers K.,  
RA Heffron J.J.A., Lehan M., Heytens L., Krivosic-Horber R., Adnet P.,  
RA Ellis F.R., Monnier N., Lunardi J., McCarthy T.V.,  
RT "Detection of a novel common mutation in the ryanodine receptor gene  
RT in malignant hyperthermia: implications for diagnosis and  
RT heterogeneity studies";  
RL Hum. Mol. Genet. 3:471-476 (1994).  
RN [14]  
RP VARIANT MH ARG-2434.  
RX MEDLINE=95152512; PubMed=7849712;  
RA Keating K.E., Quane K.A., Manning B.M., Lehan M., Hartung E.,  
RA Censier K., Urwyler A., Klausnitzer M., Muller C.R., Heffron J.J.A.,  
RA McCarthy T.V.;  
RT "Detection of a novel RYR1 mutation in four malignant hyperthermia  
RT pedigrees";  
RL Hum. Mol. Genet. 3:1855-1858 (1994).  
RN [15]  
RP VARIANT MH ARG-2434.  
RX MEDLINE=95187158; PubMed=7881417;  
RA Phillips M.S., Khanna V.K., de Leon S., Frodis W., Britt B.A.,  
RA McLennan D.H.;  
RT "The substitution of Arg for Gly2433 in the human skeletal muscle  
RT ryanodine receptor is associated with malignant hyperthermia";  
RL Hum. Mol. Genet. 3:2181-2186 (1994).  
RN [16]  
RP VARIANT MH CYS-614.  
RC TISSUE=Blood;  
RX MEDLINE=95271229; PubMed=7751854;  
RA Moroni I., Gonano E.F., Comi G.P., Tegazzin V., Prella A., Bordoni A.,  
RA Bresolin N., Scarlato G., point mutation and malignant hyperthermia  
RT "Ryanodine receptor gene point mutation and malignant hyperthermia  
RT susceptibility";  
RL J. Neurol. 242:127-133 (1995).  
RN [17]

VARIANT MH ARG-35.  
 RX MEDLINE=97219028; PubMed=9066328;  
 RA Lynch P.J., Krivosic-Horber R., Reyford H., Monnier N., Quane K.A.,  
 RA Adnet P., Haudecoeur G., Krivosic I., McCarthy T.V., Lunardi J.,  
 RT "Identification of heterozygous and homozygous individuals with the  
 RT novel R191 mutation Cys55Arg in a large kindred.";  
 RL Anesthesiology 86:620-626(1997).  
 RN [18]  
 RN VARIANT MH LEU-614.  
 RX MEDLINE=98051290; PubMed=9389851;  
 RA Quane K.A., Ordling H., Keating K.E., Manning B.M., Heine R.,  
 RA Bendixen D., Berg K., Krivosic-Horber R., Lehmann-Horn F.,  
 RA Fagerlund T.H., McCarthy T.V.,  
 RT "Detection of a novel mutation in the ryanodine receptor gene in the  
 RT ryanodine receptor in malignant hyperthermia.";  
 RL Br. J. Anaesth. 79:332-337(1997).  
 RN [19]  
 RN VARIANT MH TRP-552.  
 RX MEDLINE=97284075; PubMed=9138151;  
 RA Keating K.E., Giblin L., Lynch P.J., Quane K.A., Lehane M.,  
 RA Heffron J.J.A., McCarthy T.V.,  
 RT "Detection of a novel mutation in the ryanodine receptor gene in an  
 RT Irish malignant hyperthermia pedigree: correlation of the IVCT  
 RT response with the affected and unaffected haplotypes.";  
 RL J. Med. Genet. 34:291-296(1997).  
 RN [20]  
 RN VARIANTS MH CYS-2163; MET-2168 AND MET-2206, AND VARIANT OCD/MH  
 RP HIS-2163.  
 RX MEDLINE=98163444; PubMed=9497245;  
 RA Manning B.M., Quane K.A., Ordling H., Urwyler A., Tegazzin V.,  
 RA Lehane M., O'Halloran J., Hartung E., Giblin L.M., Lynch P.J.,  
 RA Vaughan P., Censier K., Bendixen D., Comi G.P., Heytens L.,  
 RA Monsieus K., Fagerlund T.H., Wolz W., Heffron J.J.A., Mueller C.R.,  
 RA McCarthy T.V.,  
 RT "Identification of novel mutations in the ryanodine-receptor gene  
 RT (RYR1) in malignant hyperthermia: genotype-phenotype correlation.";  
 RL Am. J. Hum. Genet. 62:599-609(1998).  
 RN [21]  
 RN VARIANTS MH CYS-2459 AND HIS-2459.  
 RX MEDLINE=9811378; PubMed=9450902;  
 RA Manning B.M., Quane K.A., Lynch P.J., Urwyler A., Tegazzin V.,  
 RA Krivosic-Horber R., Censier K., Comi G.P., Adnet P., Wolz W.,  
 RA Lunardi J., Muller C.R., McCarthy T.V.,  
 RT "Query Match 75.0%; Score 24; DB 1; Length 5038;  
 RT Best Local Similarity 50.0%; Pred. No. 4.1e+02;  
 RT Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 OY 3 GVXXXIRP 10  
 DB 1766 GVTSLRP 1773  
 RESULT 8  
 YJ53\_AQUAE STANDARD; PRT; 126 AA.  
 AC 067766;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein AQ\_1953.  
 GN AQ\_1953.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
 OC NCBI\_TaxID=63363;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Shead M.A., Keller M., Auway M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus.";  
 RL Nature 392:353-358(1998).  
 CC -1- SIMILARITY: Belongs to the rnsA family.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AB000765; AAC07739.1; ALT\_INIT.  
 DR InterPro; IPR008822; Rusa.  
 DR Pfam; PF05866; Rusa; 1.  
 KW Hypothetical protein; Hydrolase; Nuclease; Endonuclease;  
 KW Complete proteome.  
 SQ SEQUENCE 126 AA; 14778 MW; 05BC64FAA7B9F9B1 CRC64;  
 Query Match 71.9%; Score 23; DB 1; Length 126;  
 Best Local Similarity 62.5%; Pred. No. 24;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 3 GVXXXIRP 10  
 DB 115 GVIIKRP 122  
 RESULT 9  
 ORN\_STRAW STANDARD; PRT; 200 AA.  
 ID ORN\_STRAW  
 AC Q82CS8;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Oligoribonuclease [EC 3.1.-.-].  
 GN ORN OR ORNA OR SAV5260.  
 OS Streptomyces avermitilis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=33903;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MA-4580 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=2147403; PubMed=11572948;  
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,  
 RT "Genome sequence of an industrial microorganism Streptomyces  
 RT avermitilis: deducing the ability of producing secondary  
 RT metabolites.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MA-4580 / ATCC 31267 / NCIMB 12804 / NRRL 8165;  
 RX MEDLINE=22608306; PubMed=12692562;  
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,  
 RA Sakaki Y., Hattori M., Omura S.,  
 RT "Complete genome sequence and comparative analysis of the industrial  
 RT microorganism Streptomyces avermitilis.";  
 RL Nat. Biotechnol. 21:526-531(2003)  
 CC -1- FUNCTION: 3'-to-5' exoribonuclease specific for small  
 CC oligoribonucleotides (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: Belongs to the oligoribonuclease family.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).

```
CC EMBL; AP005042; BAC72972.1; -.
DR HAMAP; MF 00045; -.
DR InterPro; IPR006055; Exonuclease.
DR Pfam; PF00929; Exonuclease; 1.
DR SMART; SM00479; EXOIII; 1.
KW Hydrolase; Exonuclease; Nuclease; Complete proteome.
FT ACT_SITE 126 126 POTENTIAL.
SQ SEQUENCE 200 AA; 21995 MW; 41D2864E2F2915E6 CRC64;

Query Match
Best Local Similarity 71.9%; Score 23; DB 1; Length 200;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 40 GVDIVIRP 47

RESULT 10
ID ORN_STRCO STANDARD; PRT; 200 AA.
AC P57666; Q9KWH7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oligoribonuclease (EC 3.1.-.-).
GN ORN OR ORNA OR SCO2793 OR 2SCC13.01.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=A3(2) / M130;
RX MEDLINE=20372663; PubMed=10913103;
RA Ohnishi Y., Nishiyama Y., Sato R., Kameyama S., Horinouchi S.;
RT "An oligoribonuclease gene in Streptomyces griseus.";
RL J. Bacteriol. 182:4647-4653(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphry L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- FUNCTION: 3'-to-5' exoribonuclease specific for small
CC oligoribonucleotides.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the oligoribonuclease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB036424; BAB03461.1; -.
CC EMBL; AL939114; CAC10102.1; -.
CC HAMAP; MF 00045; -.
CC InterPro; IPR006055; Exonuclease.
CC Pfam; PF00929; Exonuclease; 1.
```

```
DR SMART; SM00479; EXOIII; 1.
KW Hydrolase; Exonuclease; Nuclease; Complete proteome.
FT ACT_SITE 126 126 POTENTIAL.
SQ SEQUENCE 200 AA; 22022 MW; 9E8E89B0995EF577 CRC64;

Query Match
Best Local Similarity 71.9%; Score 23; DB 1; Length 200;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 40 GVDIVIRP 47

RESULT 11
ID ORN_STRGR STANDARD; PRT; 201 AA.
AC P57667; Q9S165;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oligoribonuclease (EC 3.1.-.-).
GN ORN OR ORNA.
OS Streptomyces griseus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 13350;
RX MEDLINE=20010118; PubMed=10540289;
RA Ohnishi Y., Kameyama S., Onaka H., Horinouchi S.;
RT "The A-factor regulatory cascade leading to streptomycin biosynthesis
RT in Streptomyces griseus: identification of a target gene of the
RT A-factor receptor.";
RL Mol. Microbiol. 34:102-111(1999).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=IFO 13350;
RX MEDLINE=20372653; PubMed=10913103;
RA Ohnishi Y., Nishiyama Y., Sato R., Kameyama S., Horinouchi S.;
RT "An oligoribonuclease gene in Streptomyces griseus.";
RL J. Bacteriol. 182:4647-4653(2000).
CC -1- FUNCTION: 3'-to-5' exoribonuclease specific for small
CC oligoribonucleotides.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -1- SIMILARITY: Belongs to the oligoribonuclease family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB023785; BAA86266.1; -.
CC PIR; T44489; T44489.
CC HAMAP; MF 00045; -.
CC InterPro; IPR006055; Exonuclease.
CC Pfam; PF00929; Exonuclease; 1.
CC SMART; SM00479; EXOIII; 1.
KW Hydrolase; Exonuclease; Nuclease.
FT ACT_SITE 126 126 POTENTIAL.
SQ SEQUENCE 201 AA; 22149 MW; 9436B4DA313B2D65 CRC64;

Query Match
Best Local Similarity 71.9%; Score 23; DB 1; Length 201;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 40 GVDIVIRP 47
```

```

RESULT 12
VATD THEAC
ID VATD_THEAC STANDARD; PRT; 213 AA.
AC Q9H6E3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE V-type ATP synthase subunit D (EC 3.6.3.14) (V-type ATPase subunit D).
GN ATPD OR TA0006.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Friesman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton
CC gradient across the membrane.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SIMILARITY: Belongs to the V-ATPase D subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL445063; CAC11155.1;
CC HAMAP; MF_00271; -; 1.
CC InterPro; IPR002699; ATPsynt_Deb.
CC Pfam; PF01813; ATP-synt_D; 1.
CC ProDom; PD004122; ATPsynt_Dsub; 1.
CC TIGRFAMs; TIGR00309; V-ATPase_subd; 1.
CC Hydrolase; ATP synthetis; Hydrogen ion transport; Complete proteome.
CC SEQUENCE 213 AA; 24839 MW; 35254A0E089F7CB1 CRC64;
Query Match 71.9%; Score 23; DB 1; Length 213;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GVXXXIRP 10
DB 4 GYNMIRP 11
EWSL; AL445063; CAC11155.1;
HAMAP; MF_00271; -; 1.
InterPro; IPR002699; ATPsynt_Deb.
Pfam; PF01813; ATP-synt_D; 1.
ProDom; PD004122; ATPsynt_Dsub; 1.
TIGRFAMs; TIGR00309; V-ATPase_subd; 1.
Hydrolase; ATP synthetis; Hydrogen ion transport; Complete proteome.
SEQUENCE 213 AA; 24839 MW; 35254A0E089F7CB1 CRC64;
Query Match 71.9%; Score 23; DB 1; Length 213;
Best Local Similarity 62.5%; Pred. No. 39;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GVXXXIRP 10
DB 4 GYNMIRP 11
RESULT 13
T2N4 NEILA
ID T2N4 NEILA STANDARD; PRT; 243 AA.
AC P50183;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Type II restriction enzyme NlaIV (EC 3.1.21.4) (Endonuclease NlaIV)
DE (K.NlaIV).
GN NLAIV.
OS Neisseria lactamica.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=486;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 23970 / NRCC 2118;

```

```

RX MEDLINE=94247353; PubMed=8190068;
RA Lau P.C.K., Forghani F., Labbe D., Bergeron H., Brousseau R.,
RA Holtke H.J.;
RT "The NlaIV restriction and modification genes of Neisseria lactamica
RT are flanked by leucine biosynthesis genes.";
RL Mol. Gen. Genet. 243:24-31 (1994).
RN [2]
ERRATUM.
RP Lau P.C.K., Forghani F., Labbe D., Bergeron H., Brousseau R.,
RA Holtke H.J.;
RL Mol. Gen. Genet. 244:167-167 (1994).
CC -!- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE GGNCC AND
CC CLEAVES AFTER N-3.
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC specific double-stranded fragments with terminal 5'-phosphates.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U06074; AAA53238.1;
CC PIR; S43887; S43887.
CC REBASE; 1342; NlaIV.
CC Restriction system; Hydrolase; Nuclease; Endonuclease.
CC SEQUENCE 243 AA; 28826 MW; 3068E972D6736A49 CRC64;
Query Match 71.9%; Score 23; DB 1; Length 243;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GVXXXIRP 10
DB 164 GWHKIRP 171
RESULT 14
UCRI RAT
ID UCRI_RAT STANDARD; PRT; 256 AA.
AC P20788;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ubiquinol-cytochrome c reductase iron-sulfur subunit, mitochondrial
DE precursor (EC 1.10.2.2) (Rieske iron-sulfur protein) (RISP)
DE (Fragment).
DE UQCFS1.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=89165853; PubMed=2538119;
RA Nishikimi M., Hosokawa Y., Toda H., Suzuki H., Ozawa T.;
RT "Cloning and sequence analysis of a cDNA encoding the Rieske
RT iron-sulfur protein of rat mitochondrial cytochrome bcl complex.";
RL Biochem. Biophys. Res. Commun. 159:19-25 (1989).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis.
CC -!- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2
CC ferrocytochrome c.
CC -!- COFACTOR: Binds 1 2Fe-2S iron sulfur cluster per subunit (By
CC similarity).
CC -!- SUBUNIT: The bcl complex contains 11 subunits; 3 respiratory
CC subunits (cytochrome b, cytochrome c1, Rieske protein), 2 core
CC proteins and 6 low-molecular weight proteins.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.

```

CC -|- MISCELLANEOUS: The Rieske protein is a high potential 2Fe-2S  
 CC protein.  
 CC -|- SIMILARITY: Belongs to the Rieske family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M24542; AAA42051.1; --  
 CC PIR: A32296; A32296.  
 CC HSSP: FL3272; LRIE.  
 CC InterPro: IPR005805; Rieske.  
 CC InterPro: IPR005806; Rieske dom.  
 CC InterPro: IPR006317; Rieske proteo.  
 CC InterPro: IPR004192; UCR\_TM\_region.  
 CC Pfam: PF00355; Rieske; 1-  
 CC Pfam: PF02921; UCR\_TM; 1.  
 CC PRINTS: PR00162; RIESKE.  
 CC TIGRAPS: TIGR01416; Rieske\_protoso; 1.  
 CC PROSITE: PS00199; RIESKE\_1; 1.  
 CC PROSITE: PS00200; RIESKE\_2; 1.  
 CC Mitochondrion; Electron transport; Respiratory chain; Metal-binding;  
 KW Iron-sulfur; Iron; 2Fe-2S; Oxidoreductase; Inner membrane;  
 KW Transmembrane; Transit peptide.  
 FT NON\_TER 1  
 FT TRANSIT <1 60 MITOCHONDRION  
 FT CHAIN 61 256 UBIQUINOL-CYTOCHROME C REDUCTASE IRON-  
 FT SULFUR SUBUNIT.  
 FT METAL 199 199 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 201 201 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 218 218 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 221 221 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT DISULFID 204 220 BY SIMILARITY.  
 FT SEQUENCE 256 AA; 27698 MW; E44817E0D3807330 CRC64;  
 SQ  
 Query Match 71.9%; Score 23; DB 1; Length 256;  
 Best Local Similarity 50.0%; Pred. NO. 47;  
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 3 GVXXIRP 10  
 Db 3 GVAGLRP 10  
 ID UCRI\_BOVIN STANDARD; PRT; 274 AA.  
 AC P13272; P07588;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Ubiquinol-cytochrome C reductase iron-sulfur subunit, mitochondrial  
 DE Precursor (EC 1.10.2.2) (Rieske iron-sulfur protein) (Risp) [Contains:  
 DE Ubiquinol-cytochrome C reductase 8 kDa protein (Complex III subunit  
 DE IX)].  
 DE GN UQCRFS1.  
 OS Bos taurus (Bovine).  
 GN Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=90211231; PubMed=2157409;  
 RA Usui S., Yu L., Yu C.-A.;  
 RT "Cloning and sequencing of a cDNA encoding the Rieske iron-sulfur  
 RT protein of bovine heart mitochondrial ubiquinol-cytochrome c  
 RT reductase.";

RL Biochem. Biophys. Res. Commun. 167:575-579(1990).  
 RN [2]  
 RC REVISIONS: SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=93231976; PubMed=8386158;  
 RA Brandt U., Yu L., Yu C.-A., Trumpower B.L.;  
 RT "The mitochondrial targeting presence of the Rieske iron-sulfur  
 RT protein is processed in a single step after insertion into the  
 RT cytochrome bcl complex in mammals and retained as a subunit in the  
 RT complex.";  
 RL J. Biol. Chem. 268:8387-8390(1993).  
 RN [3]  
 RN SEQUENCE OF 79-274.  
 RP MEDLINE=87247298; PubMed=3036596;  
 RX Schaeffer H., Borchart U., Machleidt W., Link T.A., von Jagow G.;  
 RA "Isolation and amino acid sequence of the 'Rieske' iron sulfur  
 RA protein of beef heart ubiquinol:cytochrome c reductase.";  
 RL FEBS Lett. 219:161-168(1987).  
 RN [4]  
 RN SEQUENCE OF 1-78.  
 RP TISSUE=Heart;  
 RX MEDLINE=86030649; PubMed=2996928;  
 RA Borchart U., Machleidt W., Schaeffer H., Link T.A., von Jagow G.;  
 RT "Isolation and amino acid sequence of the 8 kDa DCCD-binding protein  
 RT of beef heart ubiquinol:cytochrome c reductase.";  
 RL FEBS Lett. 191:125-130(1985).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 148-274.  
 RX MEDLINE=96347356; PubMed=8736555;  
 RA Iwata S., Saynovits M., Link T.A., Michel H.;  
 RT "Structure of a water soluble fragment of the 'Rieske' iron-sulfur  
 RT protein of the bovine heart mitochondrial cytochrome bcl complex  
 RT determined by MAD phasing at 1.5-A resolution.";  
 RL Structure 4:567-579(1996).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).  
 RX MEDLINE=97349328; PubMed=9204897;  
 RA Xia D., Yu C.A., Kim H., Xia J.Z., Kachurin A.M., Zhang L., Yu L.,  
 RA Deisenhofer J.;  
 RT "Crystal structure of the cytochrome bcl complex from bovine heart  
 RT mitochondria.";  
 RL Science 277:60-66(1997).  
 RN [7]  
 RP ERRATUM.  
 RA Xia D., Yu C.A., Kim H., Xia J.Z., Kachurin A.M., Zhang L., Yu L.,  
 RA Deisenhofer J.;  
 RL Science 278:2037-2037(1997).  
 RN [8]  
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).  
 RX MEDLINE=98316377; PubMed=9651245;  
 RA Iwata S., Lee J.W., Okada K., Lee J.K., Iwata M., Rasmussen B.,  
 RA Link T.A., Ramaswamy S., Jap B.K.;  
 RT "Complete structure of the 11-subunit bovine mitochondrial cytochrome  
 RT bcl complex.";  
 RL Science 281:64-71(1998).  
 CC -|- FUNCTION: Component of the ubiquinol-cytochrome c reductase  
 CC complex (complex III or cytochrome b-c1 complex), which is a  
 CC respiratory chain that generates an electrochemical potential  
 CC coupled to ATP synthesis.  
 CC -|- FUNCTION: The transit peptide of the Rieske protein seems to form  
 CC part of the bcl complex in beef and is considered to be the  
 CC subunit IX of that complex.  
 CC -|- CATALYTIC ACTIVITY: QH(2) + 2 ferricytochrome c = Q + 2  
 CC ferrocycytochrome c  
 CC -|- COFACTOR: Binds 1 2Fe-2S iron sulfur cluster per subunit.  
 CC -|- SUBUNIT: The bcl complex contains 11 subunits; 3 respiratory  
 CC subunits (cytochrome b, cytochrome c1, Rieske protein), 2 core  
 CC proteins and 6 low-molecular weight proteins.  
 CC -|- SUBCELLULAR LOCATION: Mitochondrial inner membrane.  
 CC -|- MISCELLANEOUS: The Rieske protein is a high potential 2Fe-2S  
 CC protein.  
 CC -|- SIMILARITY: Belongs to the Rieske family.  
 CC -|- CAUTION: Ref.1 sequence differs from that shown by extensive

frameshifts and probable sequencing errors in the region of the  
transit peptide (1-78).

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by, and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

-----

EMBL; S58789; AAB26197.1; .  
EMBL; M34336; AAA30515.1; ALT\_FRAME.

PIR; A46063; A34660.

PDB; 1RIE; 07-DEC-96.

PDB; 1QCR; 14-OCT-98.

PDB; 1BGI; 23-FEB-99.

PDB; 1BE3; 16-FEB-99.

InterPro; IPR005805; Rieske.

InterPro; IPR005806; Rieske dom.

InterPro; IPR006317; Rieske proteo.

InterPro; IPR004192; UCR\_TM\_region.

Pfam; PF00355; Rieske; 1.

Pfam; PF02921; UCR\_TM; 1.

PRINTS; PR00162; RIESKE.

TIGRFAMs; TIGR01416; Rieske\_proteo; 1.

PROSITE; PS00199; RIESKE\_1; 1.

PROSITE; PS00200; RIESKE\_2; 1.

KW Mitochondrion; Electron transport; Respiratory chain; Metal-binding;

KW Iron-sulfur; Iron; 2Fe-2S; Oxidoreductase; Inner membrane;

KW Transmembrane; Transit peptide; 3D-structure.

CHAIN 1 78 UBIQUINOL-CYTOCHROME C REDUCTASE 8 kDa

PROTEIN

MITOCHONDRION

UBIQUINOL-CYTOCHROME C REDUCTASE IRON-

SULFUR SUBUNIT.

IRON-SULFUR (2FE-2S).

IRON-SULFUR (2FE-2S).

IRON-SULFUR (2FE-2S).

IRON-SULFUR (2FE-2S).

S -> A (IN REF. 3).

D -> G (IN REF. 3).

CC	FT	TRANSIT	1	78	MITOCHONDRION
CC	FT	CHAIN	79	274	UBIQUINOL-CYTOCHROME C REDUCTASE IRON-
CC	FT	METAL	217	217	SULFUR SUBUNIT.
CC	FT	METAL	219	219	IRON-SULFUR (2FE-2S).
CC	FT	METAL	236	236	IRON-SULFUR (2FE-2S).
CC	FT	METAL	239	239	IRON-SULFUR (2FE-2S).
CC	FT	METAL	239	239	IRON-SULFUR (2FE-2S).
CC	FT	DISULFID	222	238	IRON-SULFUR (2FE-2S).
CC	FT	CONFLICT	150	150	
CC	FT	CONFLICT	269	269	S -> A (IN REF. 3).
CC	FT	STRAND	152	155	D -> G (IN REF. 3).
CC	FT	HELIX	156	158	
CC	FT	TURN	161	162	
CC	FT	STRAND	164	169	
CC	FT	TURN	170	171	
CC	FT	STRAND	172	178	
CC	FT	HELIX	181	188	
CC	FT	TURN	189	189	
CC	FT	HELIX	192	194	
CC	FT	HELIX	201	203	
CC	FT	TURN	204	204	
CC	FT	TURN	208	209	
CC	FT	STRAND	210	214	
CC	FT	TURN	218	220	
CC	FT	STRAND	225	226	
CC	FT	TURN	227	227	
CC	FT	TURN	229	231	
CC	FT	STRAND	232	236	
CC	FT	TURN	237	240	
CC	FT	STRAND	241	244	
CC	FT	TURN	245	246	
CC	FT	STRAND	249	251	
CC	FT	STRAND	263	265	
CC	FT	STRAND	270	274	
CC	FT	SEQUENCE	274 AA;	29547 MW;	7C5FC17D2A0DD1C9 CRC64;

Query Match 71.9%; Score 23; DB 1; Length 274;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GVXXXIRP 10  
Db 21 GVAGALRP 28

Search completed: April 1, 2004, 17:39:27  
Job time : 7.75439 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2004, 17:26:03 ; Search time 34.7368 Seconds  
(without alignments)  
99.914 Million cell updates/sec

Title: US-09-833-196-6  
Perfect score: 32  
Sequence: 1 XXGVXXIRXP 11.

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues 1017041  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	81.2	366	12 Q91TT6	Q91tt6 tupaia herp
2	26	81.2	377	16 Q9PPS8	Q9pps8 ureaplasma
3	26	81.2	516	10 Q94D96	Q94d96 oryza sativ
4	25	78.1	96	10 Q8S0V8	Q8s0v8 oryza sativ
5	25	78.1	127	10 Q82310	Q82310 arabidopsis
6	25	78.1	139	2 Q84CK2	Q84ck2 streptomyc
7	25	78.1	156	10 Q7X6K7	Q7x6k7 oryza sativ
8	25	78.1	238	4 Q8NFS9	Q8nfs9 homo sapien
9	25	78.1	301	16 Q8R916	Q8r916 thermococ
10	25	78.1	353	17 Q8PWK7	Q8pwk7 methanosarc
11	25	78.1	482	16 Q9KGM0	Q9kgm0 bacillus ha
12	25	78.1	584	10 Q8L5Z4	Q8l5z4 arabidopsis
13	25	78.1	597	17 Q58637	Q58637 pyrococcus
14	25	78.1	2796	2 Q48926	Q48926 mycobacteri
15	24	75.0	135	2 Q88057	Q88057 rhodobacter
16	24	75.0	152	10 Q41349	Q41349 lycopersico

Q8ua14 agrobacteri  
Q81zk3 tropheryma  
Q81zg9 tropheryma  
Q8nj75 agaricus bi  
Q8l3f1 oryza sativ  
Q9fz96 streptomyc  
Q8k3a2 mus musculu  
Q92gw3 rickettsia  
Q859k9 staphylococ  
Q85917 staphylococ  
Q8f0d3 leptospira  
Q8u3v5 pyrococcus  
Q92au2 streptomyc  
Q9jy98 neisseria m  
Q88602 pseudomonas  
Q9zfn9 klebsiella  
Q9jtg6 neisseria m  
Q9cd55 mycobacteri  
Q53583 mycobacteri  
Q7tvm6 mycobacteri  
Q8y3z3 ralsstonia s  
Q9l122 streptomyc  
Q82hk0 streptomyc  
Q8nlf3 corynebacte  
Q9uyv7 pyrococcus  
Q8ejv2 shewanella  
Q98nn3 rhizobium l  
Q58201 pyrococcus  
Q89pi8 bradyrhizob

ALIGNMENTS

RESULT 1  
Q91TT6 PRELIMINARY; PRT; 366 AA.  
ID Q91TT6  
AC Q91TT6  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DS T22.3.  
OS Tupaia herpesvirus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae.  
OX NCBI\_TaxID=10397;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RX MEDLINE=21211637; PubMed=11312357;  
RA Bahr U., Darai G.;  
RT "Analysis and Characterization of the Complete Genome of Tupaia (Tree Shrew) Herpesvirus".  
RL J. Virol. 75:4854-4870(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2;  
RA Darai G., Bahr U.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL, AF281817; AAK57051.1; 757ED66A38D827FE CRC64;  
SQ SEQUENCE 366 AA; 42795 MW; 757ED66A38D827FE CRC64;  
Query Match 81.2%; Score 26; DB 12; Length 366;  
Best Local Similarity 62.5%; Pred. No. 87;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GYXXIRP 10

Db 162 GVSTSRP 169

RESULT 2  
Q9PPS8

```

RESULT 4
Q880Y8
ID Q880Y8 PRELIMINARY; PRT; 96 AA.
AC AC
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE F0485B12.29 protein (F0031D02.2 protein).
DE F0485B12.29 OR F0031D02.2.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
[1]
NR N
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RC "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
RC clone:P0485B12.";
RC Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
[2]
NR N
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RC Sasaki T., Matsumoto T., Yamamoto K.;
RC "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RC clone:P0031D02.";
RC Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AP003348; BAB86498.1; "-
DR EMBL; AP003230; BAB93174.1; "-
DR Gramine; Q880Y8; "-
DR SEQUENCE 96 AA; 10908 MW; 2F93BF8513818BF6 CRC64;

Query Match 78.1%; Score 25; DB 10; Length 96;
Best Local Similarity 50.0%; Pred. NO. 45;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXP 10
DD 67 GVTVSRP 74

RESULT 5
Q82310
ID Q82310 PRELIMINARY; PRT; 127 AA.
AC AC
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE AT2G25870 protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
NR N
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RC MEDLINE=30083487; PubMed=10617197;
RC Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RC Fujii C.Y., Mason T.M., Bowman C.J., Barrstead M.E., Feldblum T.V.,
RC Buell L.R., Kechum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RC Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RC Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RC Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RC Salzberg S.L., Fraser C.M., Venter J.C.;
RC "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RC thaliana.";
RC Nature 402:761-768(1999).
[2]
NR N

```



```

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003395; AAC42243.1; -.
DR PIR; G84653; G84653.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR GO; GO:0008152; P.metabolism; IEA.
DR InterPro; IPR005834; Hydrolase.
DR InterPro; IPR000150; Hypotheset_cof.
DR Pfam; PF00702; Hydrolase; 1.
DR PROSITE; PS01229; CO2; 1.
SQ SEQUENCE 127 AA; 13480 MW; E2DBBF9228D09D53 CRC64;

Query Match 78.1%; Score 25; DB 10; Length 127;
Best Local Similarity 62.5%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 19 GVSSVIRP 26
|||
|||

RESULT 6
Q84CK2 PRELIMINARY; PRT; 139 AA.
ID Q84CK2
AC Q84CK2;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE KINUL.
GN KINUL.
OS Streptomyces murayamaensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=224537;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin X.H., Mahadevan B., Grochowski L., Proteau P.J.;
RT "Molecular cloning and sequence of the kinamycin angucycline type II
RT polyketide synthase gene cluster from Streptomyces murayamaensis.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY228175; AAO65341.1; -.
DR InterPro; IPR009002; FMN binding.
SQ SEQUENCE 139 AA; 14815 MW; 55A27123A80349E4 CRC64;

Query Match 78.1%; Score 25; DB 2; Length 139;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 28 GVLTIRP 35
|||
|||

RESULT 7
Q7X6K7 PRELIMINARY; PRT; 156 AA.
ID Q7X6K7
AC Q7X6K7;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE OSJNBAC0018J19.23 protein (OSJNBAC0034E24.1 protein).
GN OSJNBAC0018J19.23 OR OSJNBAC0034E24.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.;
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,

RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC003395; AAC42243.1; -.
DR PIR; G84653; G84653.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR GO; GO:0008152; P.metabolism; IEA.
DR InterPro; IPR005834; Hydrolase.
DR InterPro; IPR000150; Hypotheset_cof.
DR Pfam; PF00702; Hydrolase; 1.
DR PROSITE; PS01229; CO2; 1.
SQ SEQUENCE 127 AA; 13480 MW; E2DBBF9228D09D53 CRC64;

Query Match 78.1%; Score 25; DB 10; Length 127;
Best Local Similarity 62.5%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 19 GVSSVIRP 26
|||
|||

RESULT 6
Q84CK2 PRELIMINARY; PRT; 139 AA.
ID Q84CK2
AC Q84CK2;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE KINUL.
GN KINUL.
OS Streptomyces murayamaensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=224537;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin X.H., Mahadevan B., Grochowski L., Proteau P.J.;
RT "Molecular cloning and sequence of the kinamycin angucycline type II
RT polyketide synthase gene cluster from Streptomyces murayamaensis.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY228175; AAO65341.1; -.
DR InterPro; IPR009002; FMN binding.
SQ SEQUENCE 139 AA; 14815 MW; 55A27123A80349E4 CRC64;

Query Match 78.1%; Score 25; DB 2; Length 139;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 28 GVLTIRP 35
|||
|||

RESULT 7
Q7X6K7 PRELIMINARY; PRT; 156 AA.
ID Q7X6K7
AC Q7X6K7;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE OSJNBAC0018J19.23 protein (OSJNBAC0034E24.1 protein).
GN OSJNBAC0018J19.23 OR OSJNBAC0034E24.1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.;
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,

RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RA J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL562956; CA804456.1; -.
DR EMBL; AL562956; CA805907.1; -.
SQ SEQUENCE 156 AA; 17088 MW; C8B47B96DFB7BD02 CRC64;

Query Match 78.1%; Score 25; DB 10; Length 156;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 30 GVTTSVRP 37
|||
|||

RESULT 8
Q8NF59 PRELIMINARY; PRT; 238 AA.
ID Q8NF59
AC Q8NF59;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE FLJ00328 protein (Fragment).
GN FLJ00328.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK090420; BAC03401.1; -.
FT NON_TER
SQ SEQUENCE 238 AA; 25220 MW; 47F56CDDC3B7A909 CRC64;

Query Match 78.1%; Score 25; DB 4; Length 238;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 GVXXXIRP 10
DB 139 GVAADIRP 146
|||
|||

RESULT 9
Q8R916 PRELIMINARY; PRT; 301 AA.
ID Q8R916
AC Q8R916;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Hypothetical protein T1E1624.
GN T1E1624.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.;
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,

```

```

RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002)
DR EMBL; AE013117; AAM24926.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 301 AA; 33042 MW; BDD505E503843B0E CRC64;

Query Match      78.1%; Score 25; DB 16; Length 301;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GVXXXIRP 10
Db 202 GVAAIIRP 209

RESULT 10
Q8PWK7 PRELIMINARY; PRT; 353 AA.
AC Q8PWK7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Molybdate ABC transporter, ATP-binding protein.
GN MML578
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=Goe1 / Goe1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12135824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AE013391; AAM31274.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR008995; MOP-like.
DR InterPro; IPR005116; TOBE.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 353 AA; 38877 MW; 26BA46955E8983D1 CRC64;

Query Match      78.1%; Score 25; DB 17; Length 353;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GVXXXIRP 10
Db 270 GVTVSIRP 277

RESULT 11
Q9KGM0 PRELIMINARY; PRT; 482 AA.
AC Q9KGM0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ATP-binding; Complete proteome.
SQ SEQUENCE 353 AA; 38877 MW; 26BA46955E8983D1 CRC64;

Query Match      78.1%; Score 25; DB 16; Length 482;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GVXXXIRP 10
Db 173 GVARSIRP 180

RESULT 12
Q8L5Z4 PRELIMINARY; PRT; 584 AA.
AC Q8L5Z4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN Ar2G25860:Ar2G25870 CR AT2G25860.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 1;
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY099804; AAM20655.1; -.
DR EMBL; AK118762; BAC43355.1; -.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006379; HAD_SF_IIB.
DR InterPro; IPR005834; Hydrolase.
DR InterPro; IPR000150; Hypothet_cof.
DR InterPro; IPR002036; UPF0054.

```

DR Pfam; PF00702; Hydrolase; 1.  
DR Pfam; PF02130; UPF0054; 1.  
DR ProDom; PD005688; UPF0054; 1.  
DR TIGRFAMs; TIGR00099; Cof-subfamily; 1.  
DR TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.  
DR TIGRFAMs; TIGR00043; TIGR00043; 1.  
DR PROSITE; PS01228; COF\_1; 1.  
DR PROSITE; PS01229; COF\_2; 1.  
DR PROSITE; PS01306; UPF0054; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 584 AA; 65305 MW; 06932128D9F2474B8 CRC64;  
  
Query Match 78.1%; Score 25; DB 10; Length 584;  
Best Local Similarity 62.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 GVXXXIRP 10  
DB 476 GVSSVIRP 483  
  
RESULT 13  
OS8537 PRELIMINARY; PRT; 597 AA.  
ID 058537  
AC 058537  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 597AA long hypothetical oligopeptide binding protein APPA.  
DN PH0807.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RA Kawaiyabashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL; AP000003; BAA29900.1; --  
DR PIR; B71130; B71130.  
DR HSP; P23847; IDPE.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000914; SBP\_bac\_5.  
DR Pfam; PF00496; SBP\_bac\_5; 1.  
KW Complete proteome.  
SQ SEQUENCE 597 AA; 67454 MW; A9F7F0CD06F7EE5C CRC64;  
  
Query Match 78.1%; Score 25; DB 17; Length 597;  
Best Local Similarity 62.5%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 GVXXXIRP 10  
DB 274 GVSVSIRP 281  
  
RESULT 14  
Q48926 PRELIMINARY; PRT; 2796 AA.  
ID 048926  
AC 048926  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Fatty acid synthase.  
GN FAS.  
OS Mycobacterium bovis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BCG-Pasteur;  
RX MEDLINE=96200863; PubMed=8621098;  
RA Fernandes N.D., Kolattukudy P.E.;  
RT "Cloning, sequencing and characterization of a fatty acid synthase-  
encoding gene from Mycobacterium tuberculosis var. Bovis BCG.";  
RL Gene 170:95-99(1996).  
DR EMBL; U36763; AAB03809.1; --  
DR PIR; JC4743; JCA743.  
DR GO; GO:0005825; C:fatty-acid synthase complex; IEA.  
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR001227; Ac trans.  
DR InterPro; IPR003965; Fatty acid synth.  
DR InterPro; IPR000794; ketoacyl\_synth.  
DR Pfam; IPR002539; Maoc\_dehydratase.  
DR Pfam; PF00698; Acyl\_transf; 1.  
DR Pfam; PF00109; ketoacyl-synt; 1.  
DR Pfam; PF02801; ketoacyl-synt C; 1.  
DR Pfam; PF01575; Maoc dehydratase; 1.  
DR PRINTS; PR01483; FASYNTHASE.  
DR PROSITE; PS00606; E\_KETOACYL\_SYNTHASE; 1.  
SQ SEQUENCE 2796 AA; 303675 MW; 47B87169126D2504 CRC64;  
  
Query Match 78.1%; Score 25; DB 2; Length 2796;  
Best Local Similarity 62.5%; Pred. No. 1.2e+03;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 GVXXXIRP 10  
DB 2704 GVHVSIRP 2711  
  
RESULT 15  
O68057 PRELIMINARY; PRT; 135 AA.  
ID 068057  
AC 068057;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacteriales;  
OC Rhodobacteriaceae; Rhodobacter.  
OX NCBI\_TaxID=1061;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SB1003;  
RX MEDLINE=97404404; PubMed=9256491;  
RA Vleck C., Paces V., Maltsev N., Paces J., Haselkorn R., Fonstein M.;  
RT "Sequence of a 189-kb segment of the chromosome of Rhodobacter  
capsulatus SB1003.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:9384-9388(1997).  
DR EMBL; AF010496; AAC16143.1; --  
DR PIR; T03490; T03490.  
KW Hypothetical protein.  
SQ SEQUENCE 135 AA; 15421 MW; B8DDDF4E69C9779 CRC64;  
  
Query Match 75.0%; Score 24; DB 2; Length 135;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 3 GVXXXIRP 10

Db 61 GVENAIRP 68

Search completed: April 1, 2004, 17:38:43  
Job time : 35.7368 secs



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 118243**

**TO: David Lukton**  
**Location: rem/3b75/3c70**  
**Art Unit: 1653**  
**Friday, April 02, 2004**

**Case Serial Number: 09/833196**

**From: Mary Jane Ruhl**  
**Location: Biotech-Chem Library**  
**Remsen 1-B55**  
**Phone: 571-272-2524**

**maryjane.ruhl@uspto.gov**

### **Search Notes**

Examiner Lukton,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl  
Technical Information Specialist  
STIC  
CM-1, Rm. 6-A-06  
605-1155



CRFE

118243

Mark Spencer  
contracted  
3/30/04 JB

SEARCH REQUEST FORM  
(STIC)

Requestor's Name: David Lukton

Examiner number: 71263

Date: 3/30/04

Art Unit: 1653

Phone number: 571-272-0952

Serial Number:

09-833 196

Mail Box: 3-C-70

Examiner Rm: 3-B-75

Results format: paper

\*\*\*\*\*

Title: Peptide antiangiogenic drugs

Applicants: HAVIV, FORTUNA; HENKIN, JACK; BRADLEY, MICHAEL  
F.; SCHNEIDER, ANDREW J.; Calvin, Douglas M

Earliest Priority Date: 5/22/98

\*\*\*\*

Please search the sequences in this case (SEQ ID NOS: 1 - 6)

1- AA 10  
2- 9  
3-  
4-  
5-  
6- 11

Sub 196  
4/1 5P

STIC  
MAR 30 2004

02

\*\*\*\*\*  
**STAFF USE ONLY**

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) _____	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: _____	Bibliographic _____	Dr.Link _____
Date Completed: _____	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems _____
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: _____	Other _____	Other (specify) _____

=> fil hcaplus  
 FILE 'HCAPLUS' ENTERED AT 16:22:58 ON 28 SEP 2003  
 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
 PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
 COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

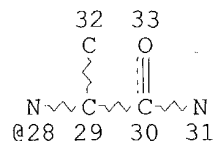
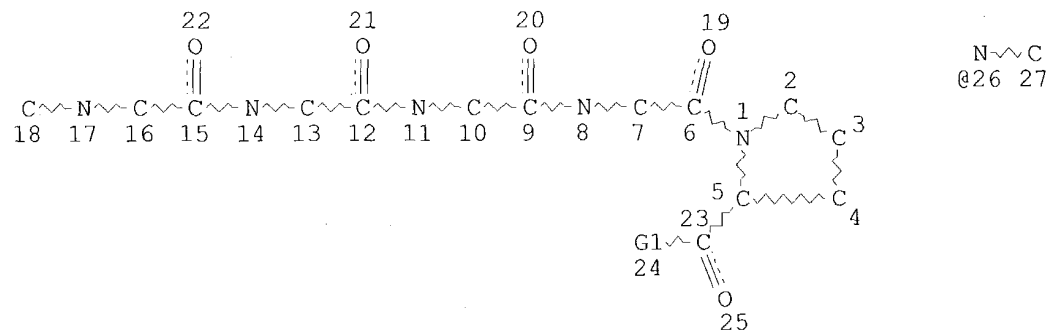
Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 28 Sep 2003 VOL 139 ISS 14  
 FILE LAST UPDATED: 26 Sep 2003 (20030926/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=>  
 =>

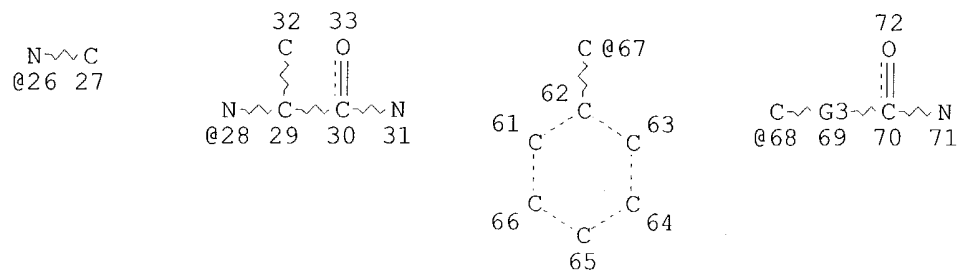
=> d stat que 126  
 L1 STR



VAR G1=26/28  
 NODE ATTRIBUTES:  
 DEFAULT MLEVEL IS ATOM  
 DEFAULT ECLEVEL IS LIMITED

GRAPH ATTRIBUTES:  
 RING(S) ARE ISOLATED OR EMBEDDED  
 NUMBER OF NODES IS 33

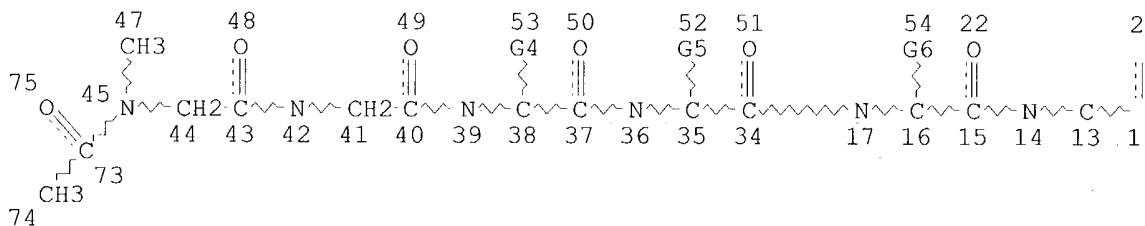
STEREO ATTRIBUTES: NONE  
 L7 121997 SEA FILE=REGISTRY SSS FUL L1  
 L24 STR



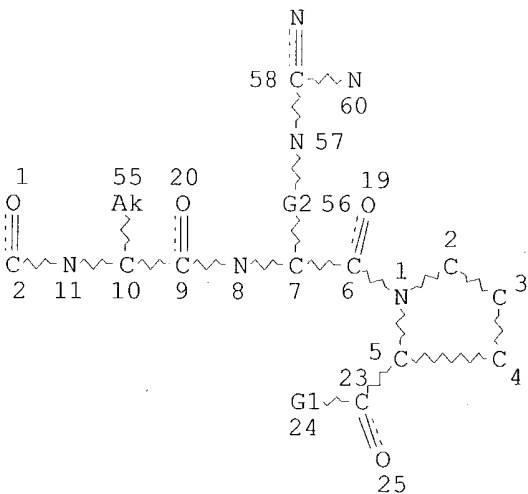
Page 1-A

59

Page 1-B



Page 2-A



Page 2-B

VAR G1=26/28

REP G2=(3-3) C

REP G3=(0-1) C

VAR G4=67/68/I-PR

VAR G5=ME/ET/I-PR/N-PR/I-BU/N-BU/T-BU/S-BU

VAR G6=76/78/82

NODE ATTRIBUTES:

DEFAULT MLEVEL IS ATOM

DEFAULT ECLEVEL IS LIMITED



GRAPH ATTRIBUTES:  
RING(S) ARE ISOLATED OR EMBEDDED  
NUMBER OF NODES IS 82

STEREO ATTRIBUTES: NONE

L25 67 SEA FILE=REGISTRY SUB=L7 SSS FUL L24  
L26 3 SEA FILE=HCAPLUS ABB=ON PLU=ON L25

=>  
=>

=> d ibib abs hitrn l26 1-3

L26 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:813873 HCAPLUS  
DOCUMENT NUMBER: 137:325642  
TITLE: Peptide antiangiogenic drugs  
INVENTOR(S): Henkin, Jack; Haviv, Fortuna; Bradley, Michael F.;  
Douglas, Calvin M.; Schneider, Andrew J.  
PATENT ASSIGNEE(S): Abbott Laboratories, USA  
SOURCE: PCT Int. Appl., 33 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002083065	A2	20021024	WO 2002-US11027	20020410
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG US 2002183242 A1 20021205 US 2001-832733 20010411 PRIORITY APPLN. INFO.: US 2001-832733 A 20010411				

OTHER SOURCE(S): MARPAT 137:325642

AB The present invention discloses peptides which are useful for inhibiting angiogenesis. Also disclosed are angiogenesis-inhibiting compns. and methods of inhibiting angiogenesis in a mammal.

IT 472967-80-9P 472967-82-1P 472967-84-3P  
472967-85-4P 472967-88-7P 472967-91-2P

RL: BSU (Biological study, unclassified); PNU (Preparation, unclassified);  
PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP  
(Preparation); USES (Uses)  
(peptide antiangiogenic drugs)

L26 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:396931 HCAPLUS  
DOCUMENT NUMBER: 135:5822  
TITLE: Preparation of N-alkylated peptides having  
antiangiogenic activity  
INVENTOR(S): Haviv, Fortuna; Henkin, Jack; Bradley, Michael F.;  
Kalvin, Douglas M.  
PATENT ASSIGNEE(S): Abbott Laboratories, USA  
SOURCE: PCT Int. Appl., 95 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001038397	A1	20010531	WO 2000-US32105	20001122
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1242455	A1	20020925	EP 2000-980685	20001122
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO, MK, CY, AL, TR			
BR 2000010934	A	20021217	BR 2000-10934	20001122
JP 2003514920	T2	20030422	JP 2001-540160	20001122
PRIORITY APPLN. INFO.:			US 1999-447099 A	19991122
			US 2000-702649 A	20001031
			WO 2000-US32105 W	20001122

OTHER SOURCE(S): MARPAT 135:5822

AB N-Alkylated peptides Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Xaa7-Xaa8-Xaa9-Xaa10-Xaa11 [Xaa3-Xaa10 represent amino acid residues (defined), at least one of which is N-alkylated; Xaa1 is absent [when Xaa2 is N-(R3)-prolyl, where R3 = C1-5alkyl], H, N-methylprolyl, or an acyl group; Xaa2 is an N-alkylated amino acid selected from N-(R3)alanyl, N-(R3)glycyl, N-(R3)norvalyl, and N-(R3)prolyl or Xaa2 is an N-unalkylated amino acid (defined); Xaa11 is OH or an amino acid amide selected from alanylamine, D-alanylamine, alanyllethylamine, D-alanyllethylamine, azaglycylamine, glycylamine, glycylethylamine, lysyl(N.epsilon.-acetyl), D-lysyl(N.epsilon.-acetyl), N-methyl-D-alanylamine, sarcosylamine, serylamine, D-serylamine, a residue NH(CH2)sCHR4R5 or NHR6 [s = 0-8, R4 = H, alkyl, 5- to 6-membered cycloalkyl; R5 = H, alkoxy, alkyl, aryl, cycloalkenyl, cycloalkyl, heterocyclyl, OH (provided that s is not zero when R5 is OH or alkoxy); R6 = H, OH]] were prep'd. for inhibiting angiogenesis. Thus, N-Ac-Sar-Gly-Val-D-Ile-Thr-NMeNva-Ile-Arg-Pro-NHEt was prep'd. by the solid phase method. Preferred compds. inhibited human endothelial cell migration by at least 51% at concns. of 10 nM or 1 nM.

IT 341012-77-9P 341012-78-0P 341012-86-0P  
 341012-87-1P 341012-95-1P 341012-96-2P  
 341012-98-4P 341012-99-5P 341013-04-5P  
 341013-05-6P 341013-44-3P 341013-45-4P  
 341013-73-8P 341013-76-1P 341014-14-0P  
 341014-36-6P 341014-38-8P 341014-57-1P  
 341014-59-3P 341014-61-7P 341014-71-9P

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (prepn. of N-alkylated peptides having antiangiogenic activity)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L26 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:764071 HCAPLUS

DOCUMENT NUMBER: 132:23191

TITLE: Preparation of peptide antiangiogenic drugs

INVENTOR(S): Henkin, Jack; Haviv, Fortuna; Bradley, Michael F.;

Kalvin, Douglas M.; Schneider, Andrew J.  
 PATENT ASSIGNEE(S): Abbott Laboratories, USA  
 SOURCE: PCT Int. Appl., 223 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9961476	A1	19991202	WO 1999-US11448	19990521
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, VZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
CA 2329250	AA	19991202	CA 1999-2329250	19990521
AU 9944075	A1	19991213	AU 1999-44075	19990521
EP 1078002	A1	20010228	EP 1999-927091	19990521
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, PT, IE, SI, FI, RO				
BR 9910639	A	20020115	BR 1999-10639	19990521
JP 2002516342	T2	20020604	JP 2000-550879	19990521
NZ 507912	A	20021025	NZ 1999-507912	19990521
NO 2000005890	A	20010112	NO 2000-5890	20001121
BG 105064	A	20010831	BG 2000-105064	20001218
PRIORITY APPLN. INFO.:				
			US 1998-83745	A 19980522
			US 1999-250574	A 19990216
			US 1999-277466	A 19990326
			WO 1999-US11448	W 19990521

OTHER SOURCE(S): MARPAT 132:23191

AB Peptides A0-A1-A2-A3-A4-A5-A6-A7-A8-A9-A10 (A0 is H or an acyl group; A10 is OH or an amino acid amide; A1-9 are amino acyl residues) or their pharmaceutically acceptable salts, esters, solvates, or prodrugs were prepd. for the treatment of angiogenesis. Thus, N-Ac-Sar-Gly-Val-D-Ile-Thr-Nva-Ile-Arg-Pro-NHEt was prepd. by the solid-phase method and assayed for in vitro angiogenic activity (87.3% at 20 nM and 76.9 at 10 nM).

IT 251581-71-2P 251581-72-3P 251581-87-0P  
 251581-88-1P 251581-89-2P 251581-90-5P  
 251583-85-4P 251583-86-5P 251583-96-7P  
 251584-10-8P 251584-11-9P 251584-17-5P  
 251584-18-6P 251584-24-4P 251584-25-5P  
 251584-26-6P 251584-27-7P 251584-28-8P  
 251584-29-9P 251584-34-6P 251584-35-7P  
 251584-36-8P 251584-37-9P 251584-38-0P  
 251584-39-1P 251584-40-4P 251584-42-6P  
 251584-43-7P 251584-44-8P 251584-45-9P  
 251584-51-7P 251584-52-8P 251584-53-9P  
 251584-54-0P 251584-55-1P 251584-81-3P  
 251584-82-4P 251584-85-7P 251585-42-9P

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (prepn. of peptide antiangiogenic drugs)

REFERENCE COUNT: 4 THERE ARE 4 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=>

=>

=> fil reg

FILE 'REGISTRY' ENTERED AT 16:23:29 ON 28 SEP 2003  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file  
provided by InfoChem.

STRUCTURE FILE UPDATES: 26 SEP 2003 HIGHEST RN 593958-55-5  
DICTIONARY FILE UPDATES: 26 SEP 2003 HIGHEST RN 593958-55-5

TSCA INFORMATION NOW CURRENT THROUGH JULY 14, 2003

Please note that search-term pricing does apply when  
conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP  
PROPERTIES for more information. See STNote 27, Searching Properties  
in the CAS Registry File, for complete details:  
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=>

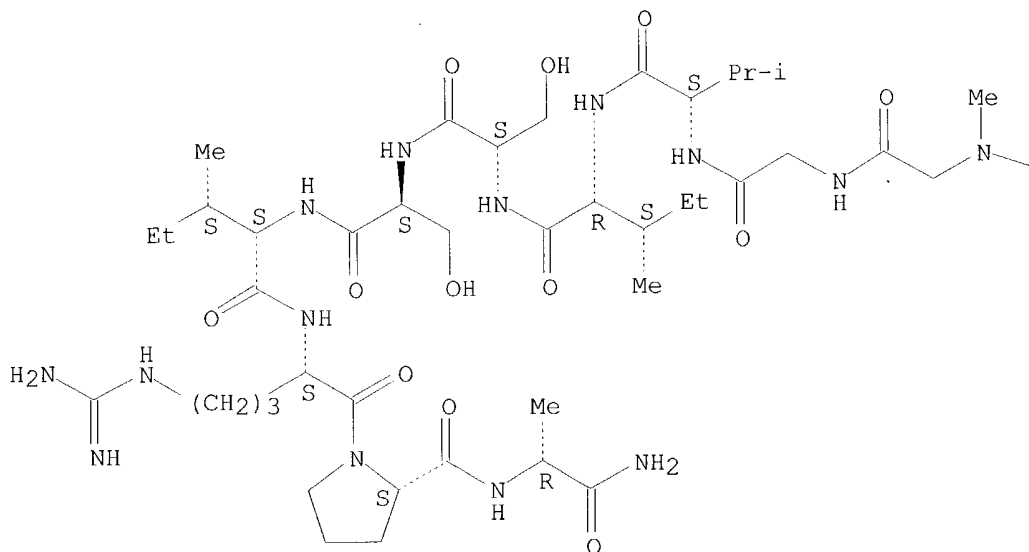
=>

=> d ide can 125 1-67

L25 ANSWER 1 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 472967-91-2 REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-  
seryl-L-seryl-L-isoleucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
OTHER NAMES:  
CN 22: PN: WO02083065 PAGE: 30 claimed protein  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C44 H78 N14 O13  
SR CA  
LC STN Files: CA, CAPLUS, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

1 Absolute stereochemistry.



PAGE 1-B

 $\text{Ac}$ 

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 137:325642

L25 ANSWER 2 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 472967-88-7 REGISTRY

472987-00-7, REBISIN  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-  
seryl-L-seryl-L-isoleucyl-L-arginyl-N-(1-methylethyl)- (9CI) (CA INDEX  
NAME)

OTHER NAMES:

CN 19: PN: WO02083065 PAGE: 30 claimed protein

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C44 H79 N13 O12

SR CA

LC STN Files: CA, CAPLUS, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



REFERENCE 1: 137:325642

RN 472967-85-4 REGISTRY

OTHER NAMES:

FS PROTEIN SEQUENCE; STEREOSEARCH

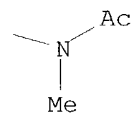
MF C51 H84 N14 O12

SR	CA
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-B



\*\*PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT\*\*

1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 137:325642

L25 ANSWER 4 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 472967-84-3 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-L-threonyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 15: PN: WO02083065 PAGE: 30 claimed protein

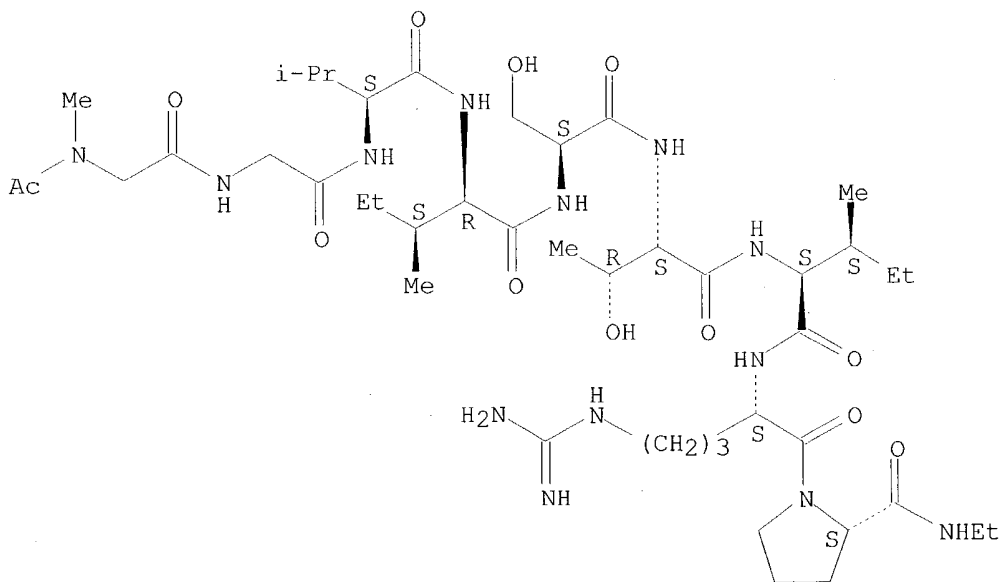
FS PROTEIN SEQUENCE; STEREOSEARCH

MF C44 H79 N13 O12

SR CA

LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



\*\*PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT\*\*

1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 137:325642

L25 ANSWER 5 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 472967-82-1 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-asparaginyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 13: PN: WO02083065 PAGE: 30 claimed protein

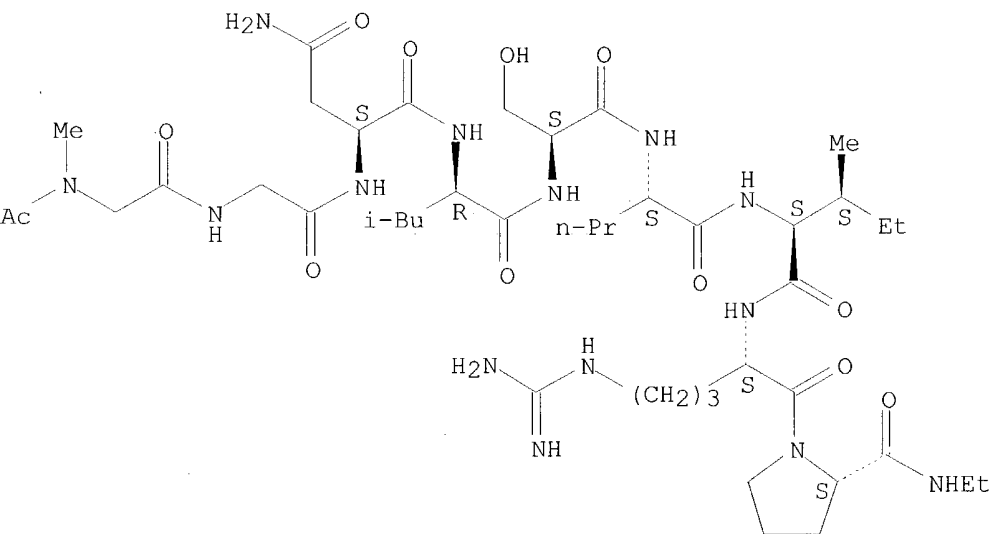
FS PROTEIN SEQUENCE; STEREOSEARCH

MF C44 H78 N14 O12

SR CA

LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.



\*\*PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT\*\*

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 137:325642

L25 ANSWER 6 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 472967-80-9 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-methionyl-L-norvalyl-L-isoleucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 11: PN: WO02083065 PAGE: 30 claimed protein

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C48 H86 N14 O11 S

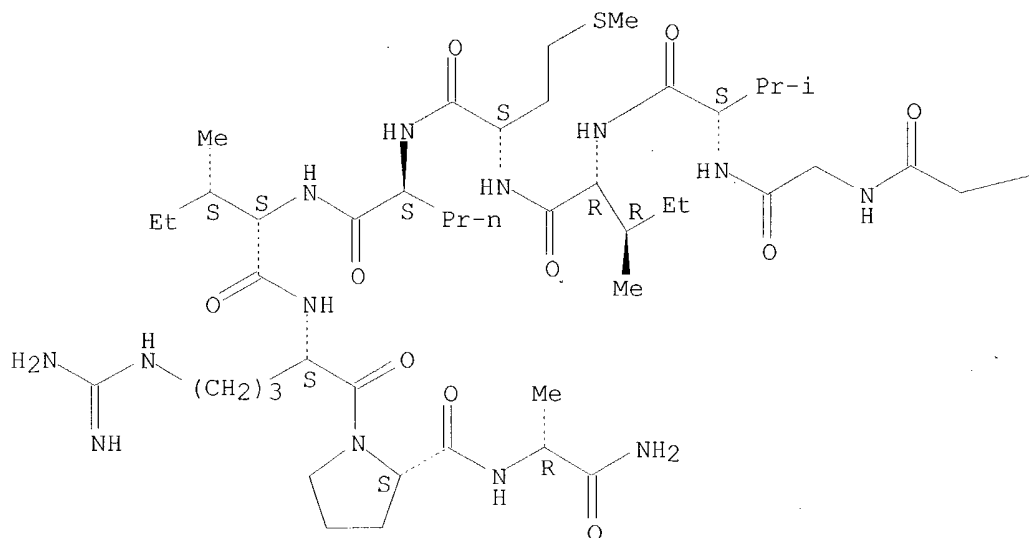
SR CA

LC STN Files: CA, CAPLUS, USPATFULL

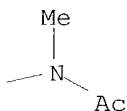
Absolute stereochemistry.



PAGE 1-A



PAGE 1-B



\*\*PROPERTY DATA AVAILABLE IN THE 'PROP' FORMAT\*\*

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

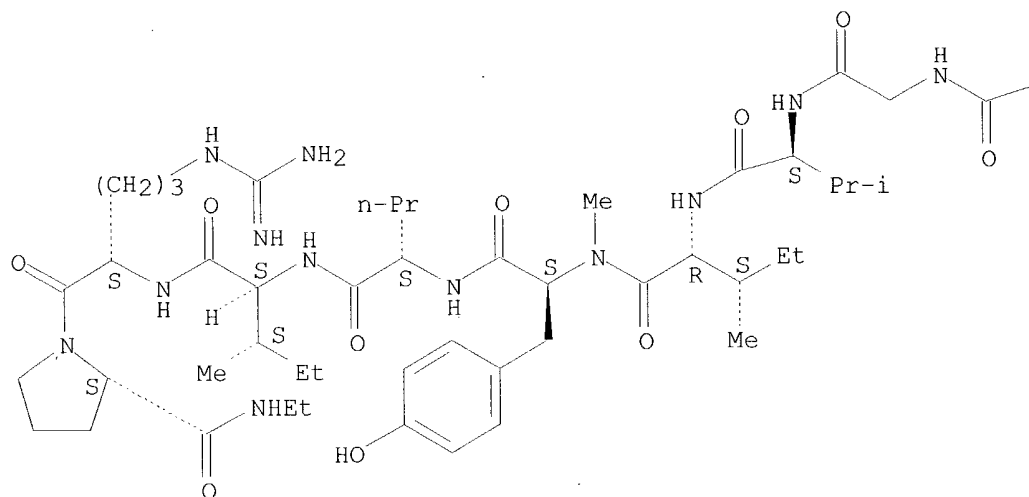
REFERENCE 1: 137:325642

L25 ANSWER 7 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341014-71-9 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-N-methyl-L-tyrosyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C52 H87 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

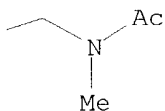
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



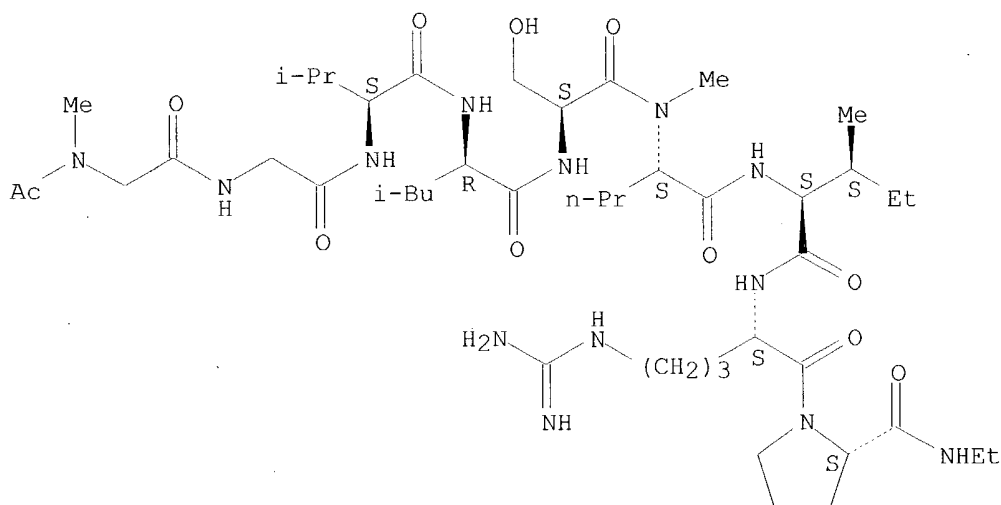
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 8 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341014-61-7 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-N-methyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



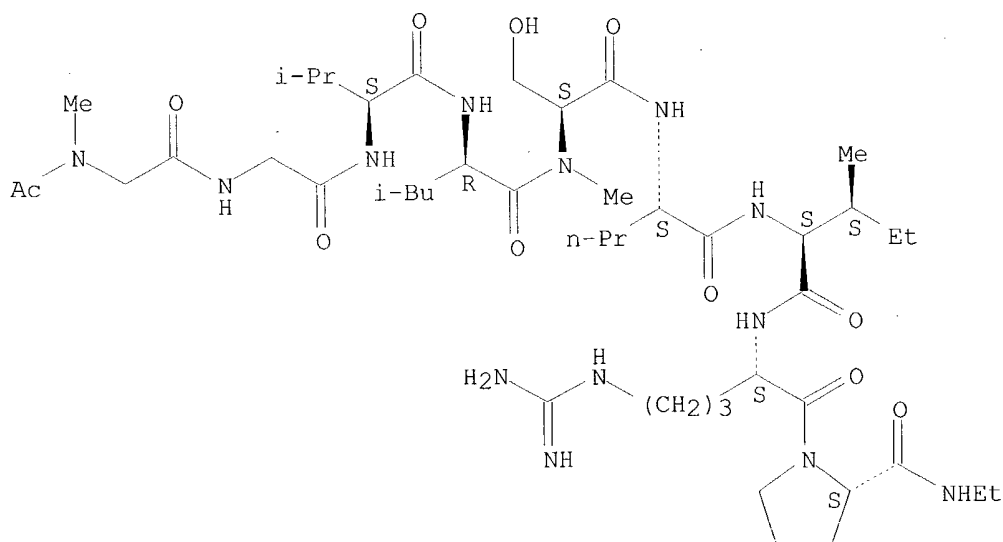
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 9 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341014-59-3 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-N-methyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C46 H83 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



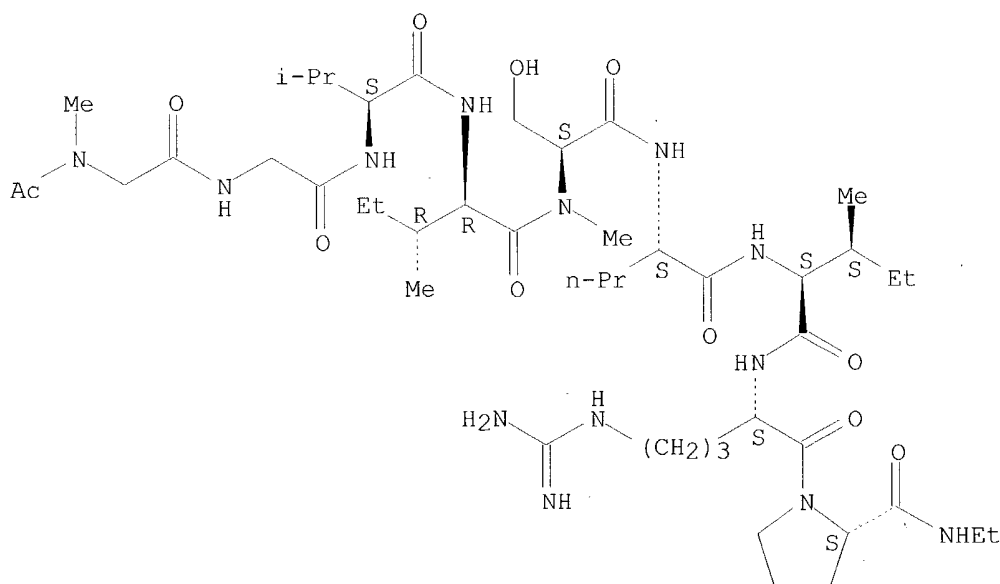
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 10 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341014-57-1 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-N-methyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

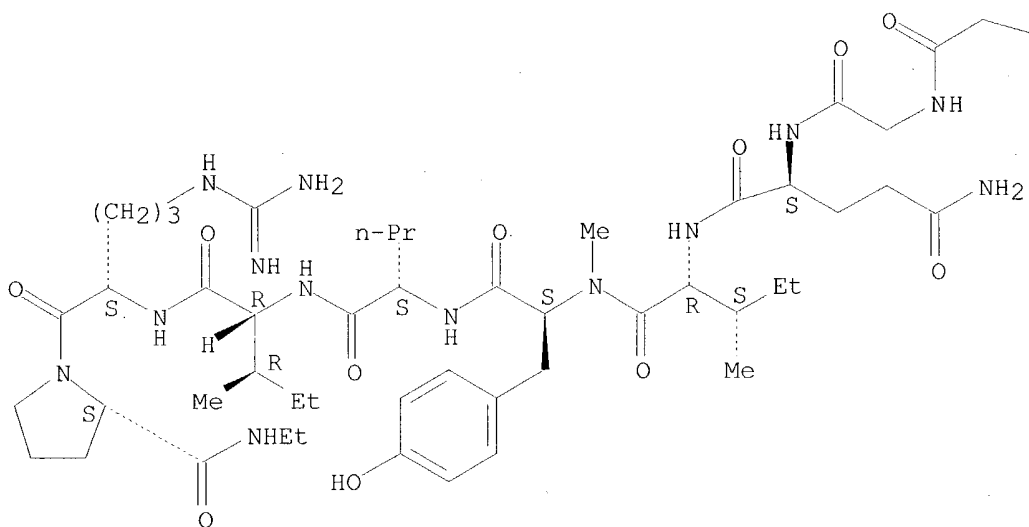
REFERENCE 1: 135:5822

L25 ANSWER 11 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341014-38-8 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-glutaminyl-D-alloisoleucyl-N-methyl-L-tyrosyl-L-norvalyl-D-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C52 H86 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

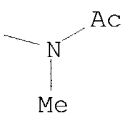
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

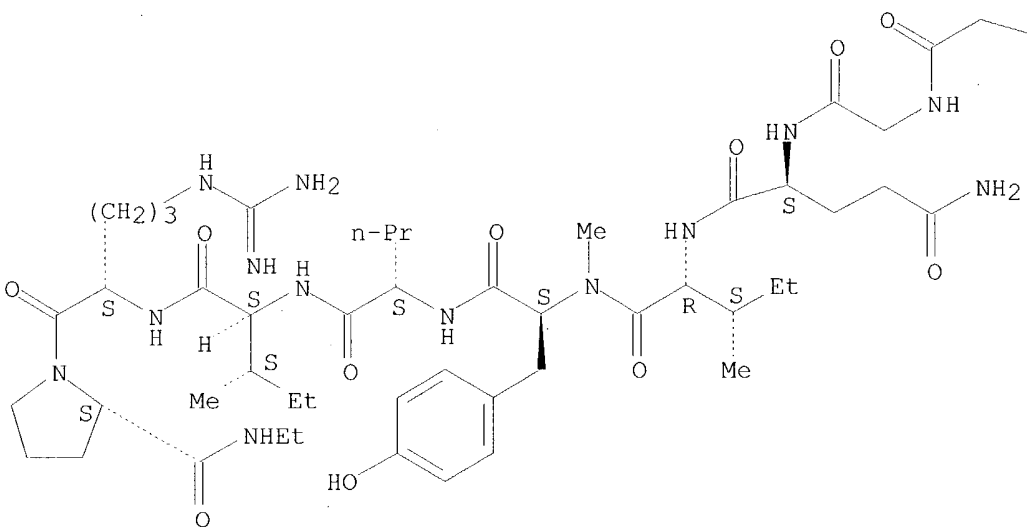
REFERENCE 1: 135:5822

L25 ANSWER 12 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341014-36-6 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-glutaminyl-D-alloisoleucyl-  
 N-methyl-L-tyrosyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA  
 INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C52 H86 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

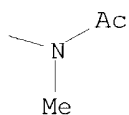
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



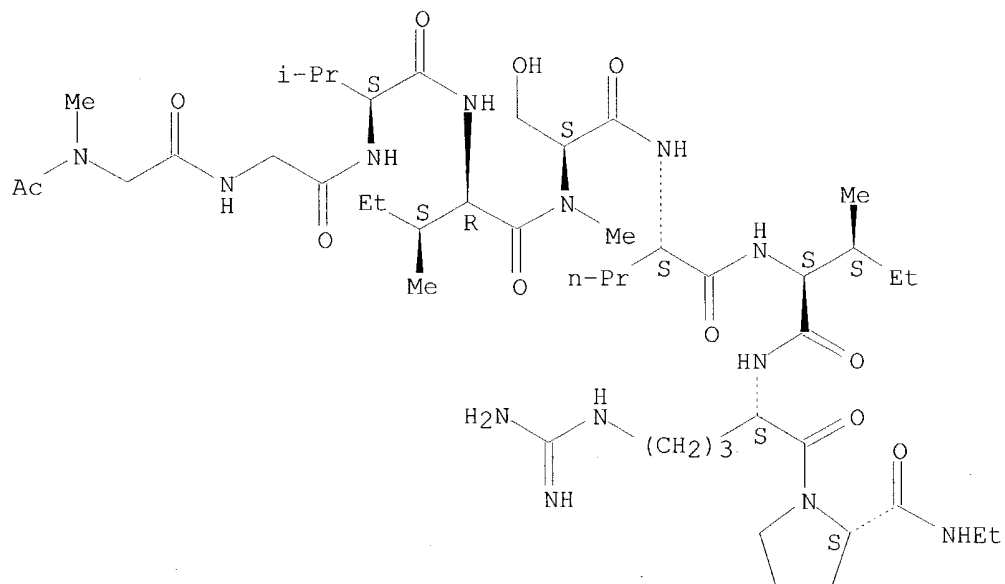
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 13 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341014-14-0 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-N-methyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

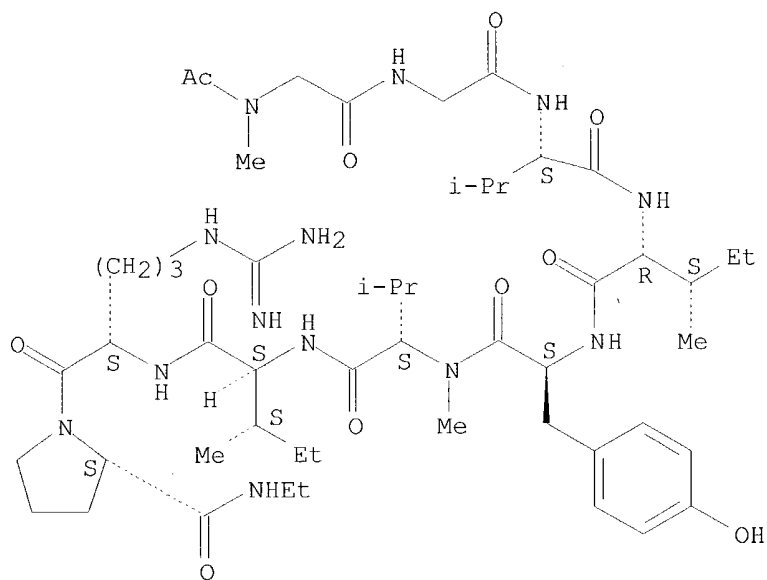


1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 14 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341013-76-1 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-tyrosyl-N-methyl-L-valyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C52 H87 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

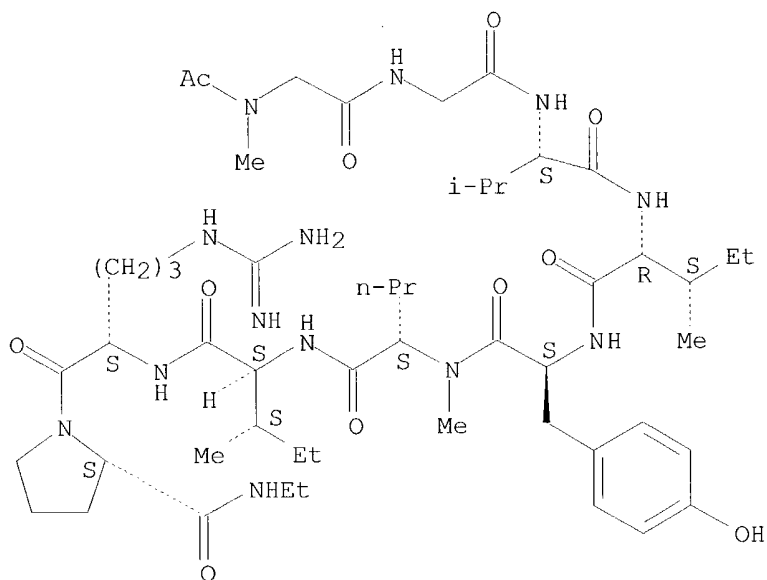
REFERENCE 1: 135:5822

L25 ANSWER 15 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341013-73-8 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-tyrosyl-N-methyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C52 H87 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.





1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 16 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341013-45-4 REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-N-methyl-L-seryl-L-isoleucyl-L-arginyl-L-prolyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H80 N14 O13 . C2 H F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

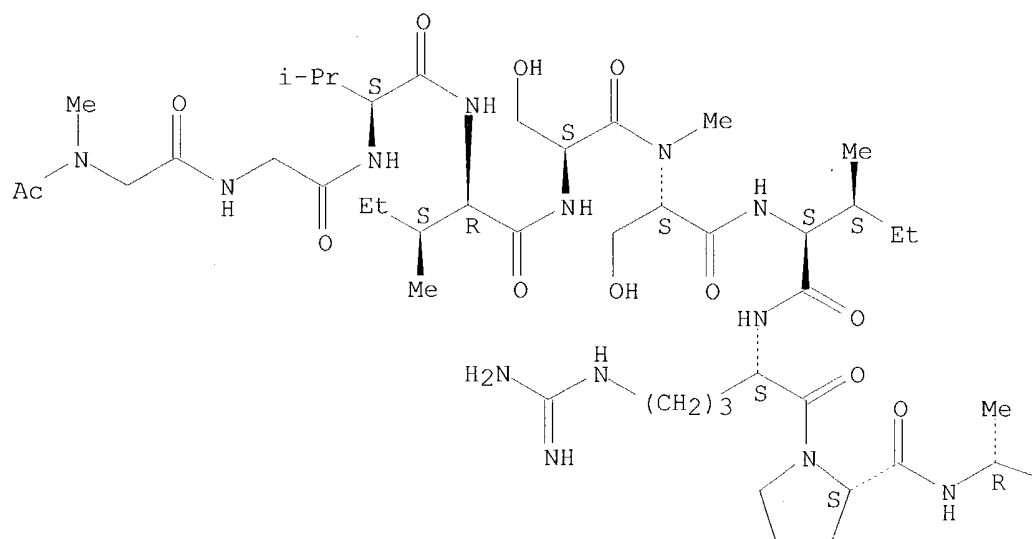
CM 1

CRN 341013-44-3  
CMF C45 H80 N14 O13

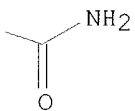
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



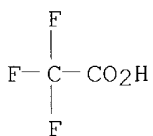
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

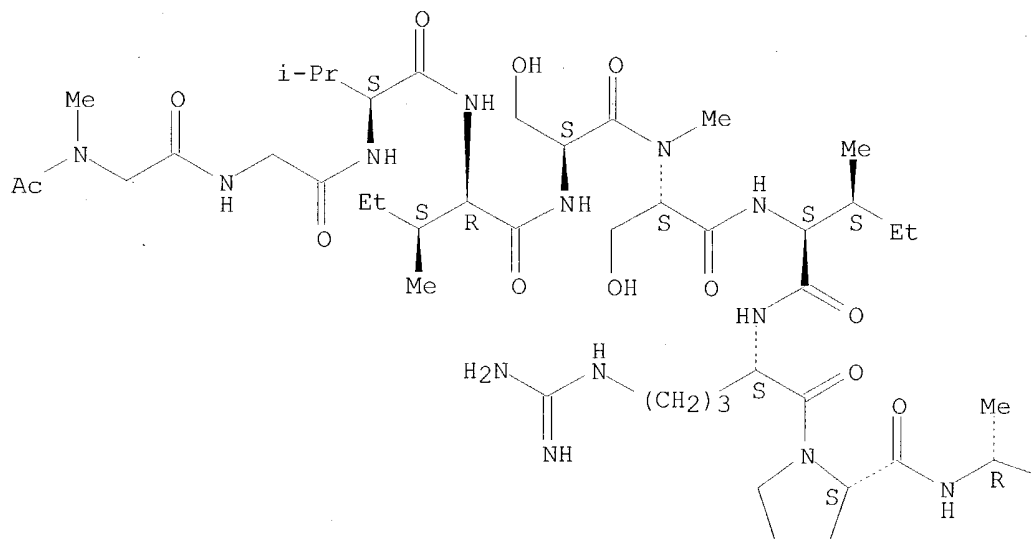
REFERENCE 1: 135:5822

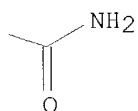
L25 ANSWER 17 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341013-44-3 REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-N-methyl-L-seryl-L-isoleucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H80 N14 O13  
CI COM  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A





1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 18 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 341013-05-6 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-asparaginyl-D-leucyl-N-methyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H80 N14 O12 . C2 H F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

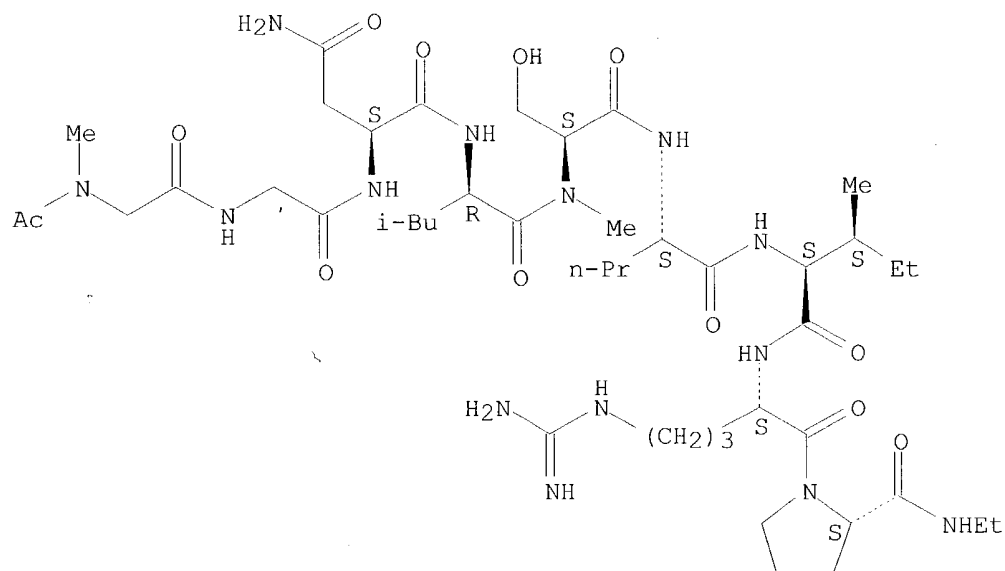
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 341013-04-5  
CMF C45 H80 N14 O12

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

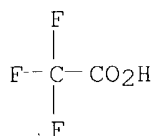
Absolute stereochemistry.



CM 2

CRN 76-05-1

CMF C2 H F3 O2



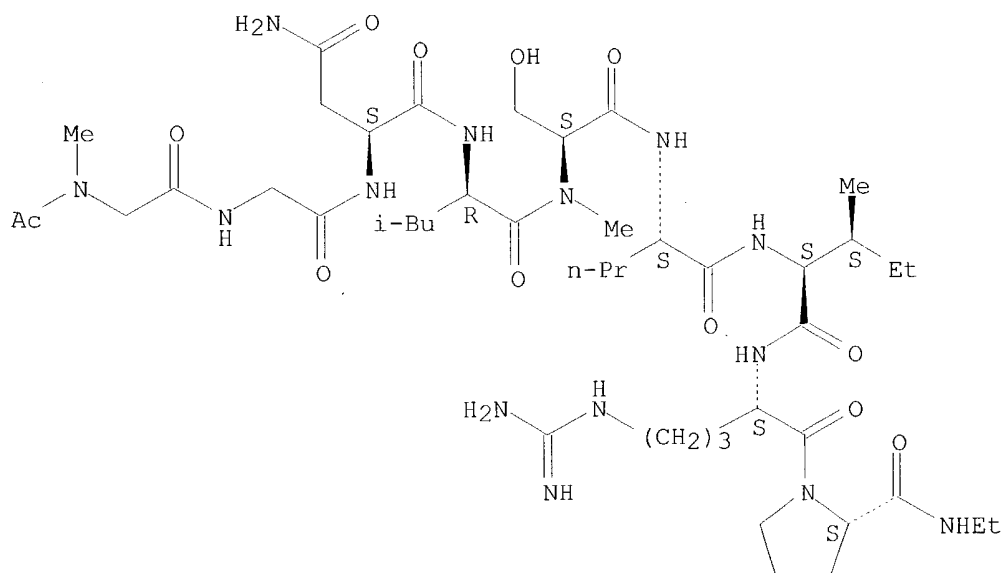
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 19 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341013-04-5 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-asparaginyl-D-leucyl-N-methyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 20 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341012-99-5 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-asparaginyl-N-methyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12 . C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

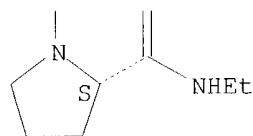
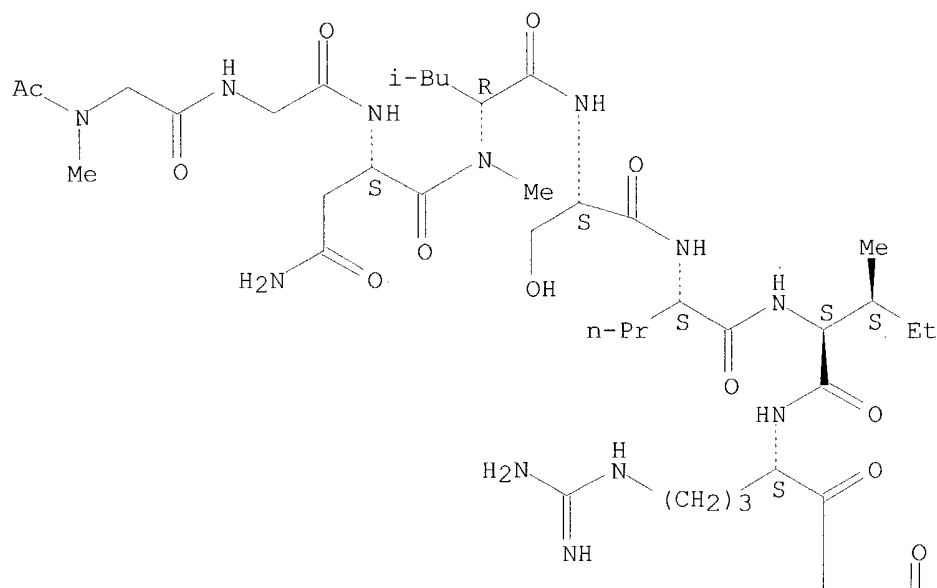
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 341012-98-4  
 CMF C45 H80 N14 O12

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

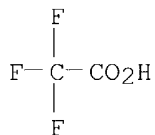
Absolute stereochemistry.



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 21 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 341012-98-4 REGISTRY

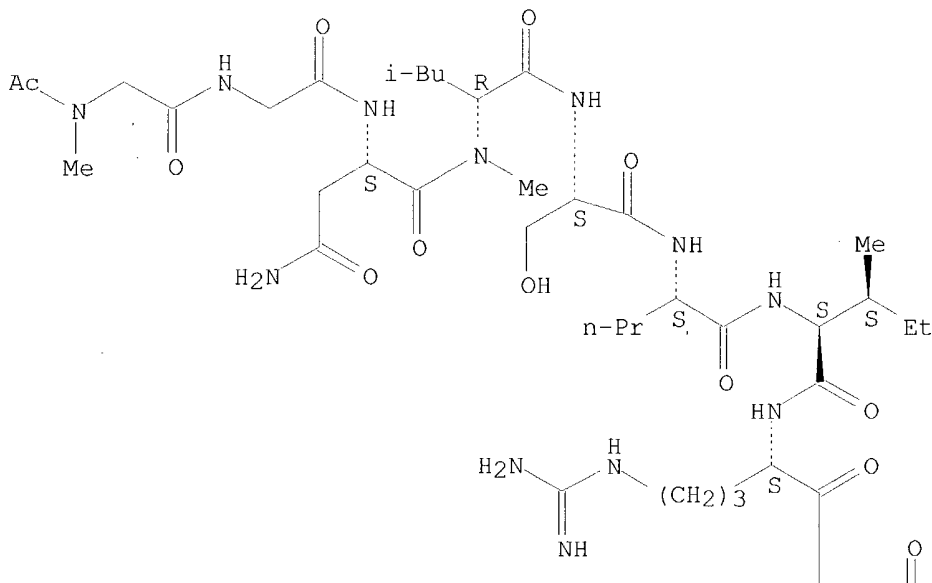
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-asparaginyl-N-methyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

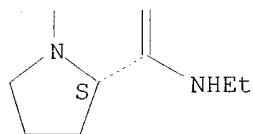
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 22 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341012-96-2 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-N-methyl-D-leucyl-L-  
 seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate)  
 (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11 . C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER



\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

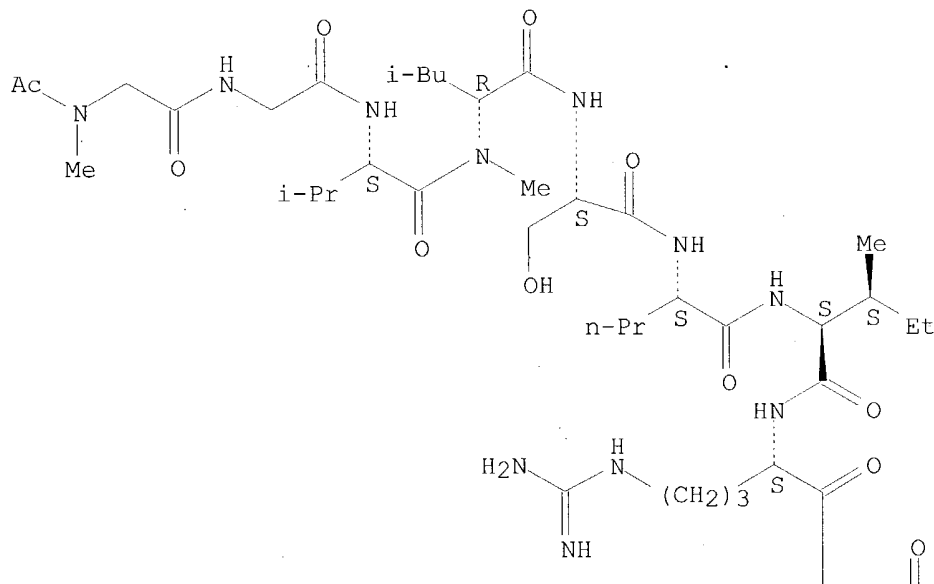
CM 1

CRN 341012-95-1  
CMF C46 H83 N13 O11

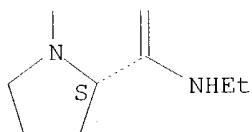
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

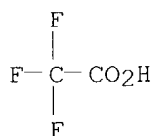


PAGE 2-A



CM 2

CRN 76-05-1  
CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

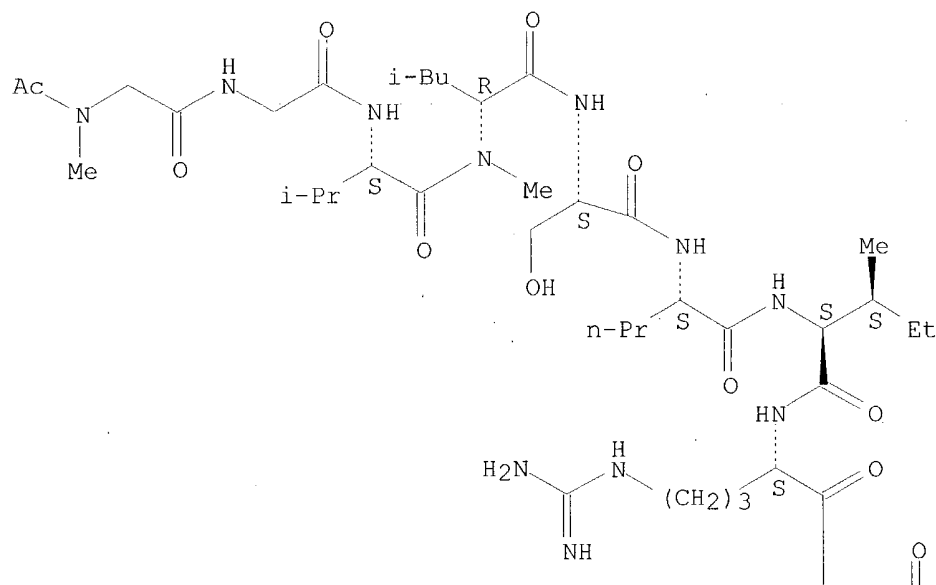
REFERENCE 1: 135:5822

L25 ANSWER 23 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341012-95-1 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-N-methyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

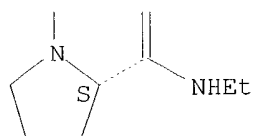
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 24 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

```

RN      341012-87-1  REGISTRY
CN      L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-N-
        methyl-L-seryl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl-,
        mono(trifluoroacetate) (salt) (9CI)  (CA INDEX NAME)
FS      PROTEIN SEQUENCE; STEREOSEARCH
MF      C44 H79 N13 O12 . C2 H F3 O2
SR      CA
LC      STN Files:  CA, CAPLUS, TOXCENTER

```

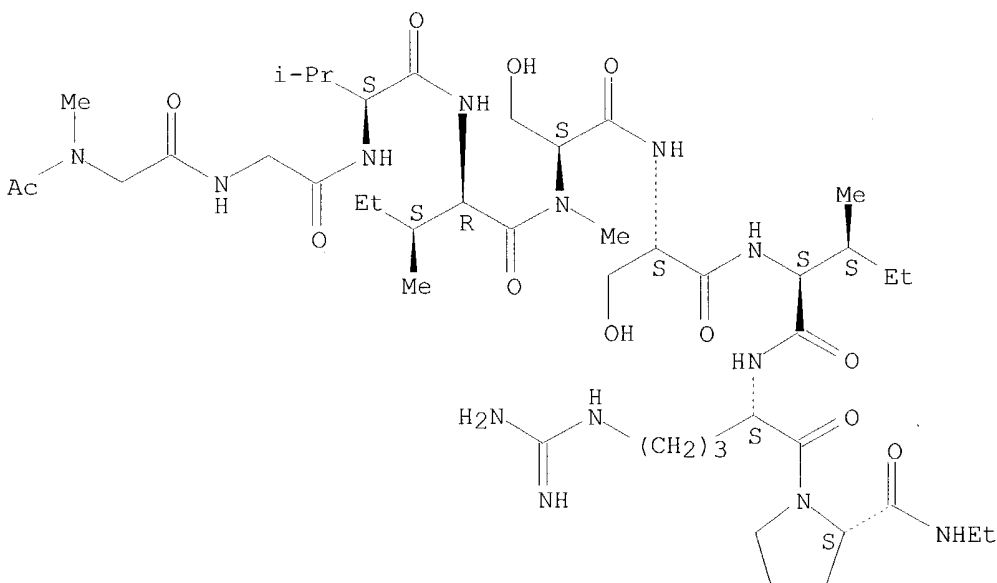
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 341012-86-0  
CMF C44 H79 N13 O12

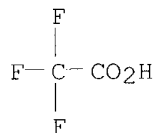
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



CM 2

CRN 76-05-1  
CMF C2 H F3 O2



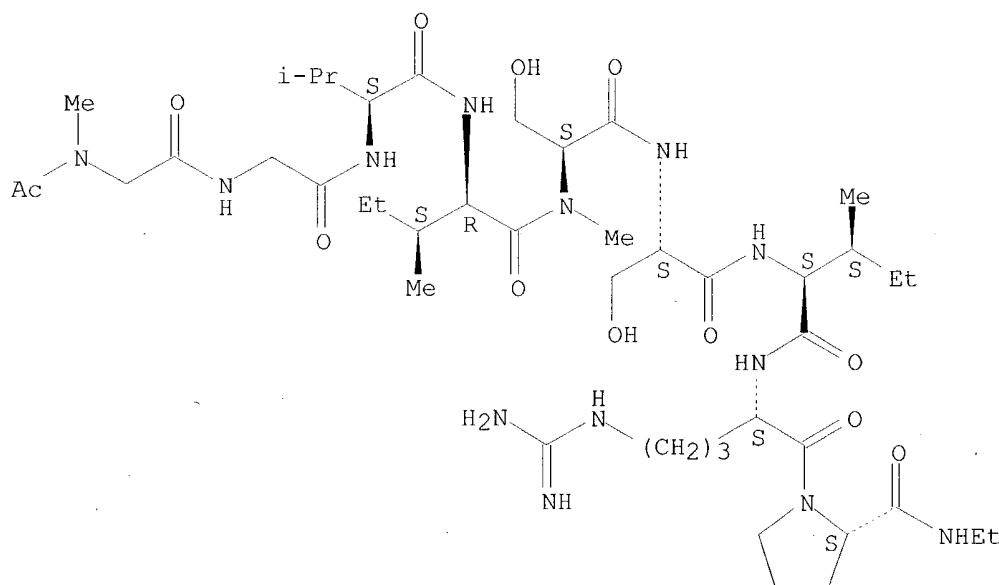
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 25 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341012-86-0 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-N-methyl-L-seryl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C44 H79 N13 O12  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 26 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 341012-78-0 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-N-methyl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C44 H79 N13 O12 . C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

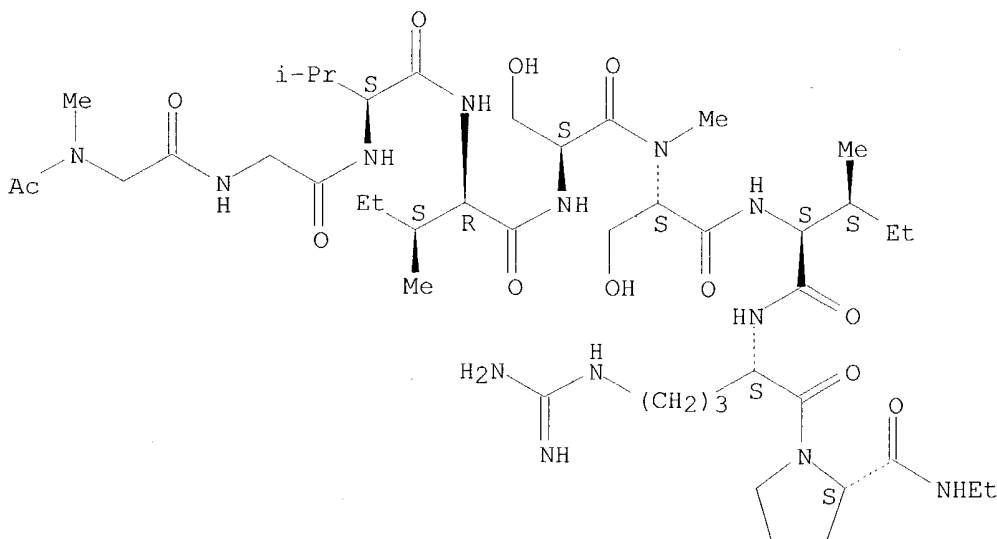
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 341012-77-9  
 CMF C44 H79 N13 O12

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

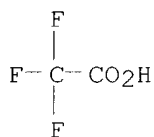
Absolute stereochemistry.



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 135:5822

L25 ANSWER 27 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 341012-77-9 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-N-methyl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C44 H79 N13 O12

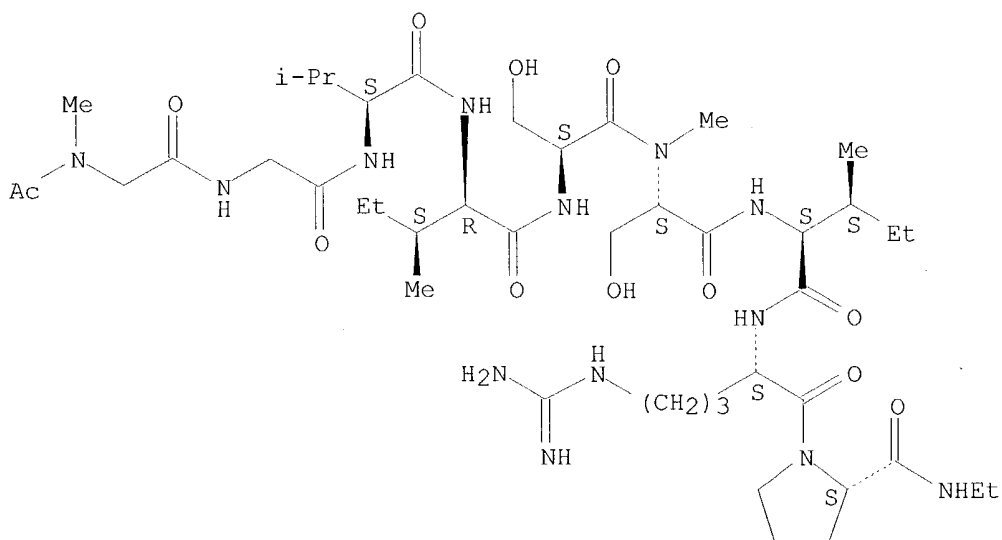
CI COM

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

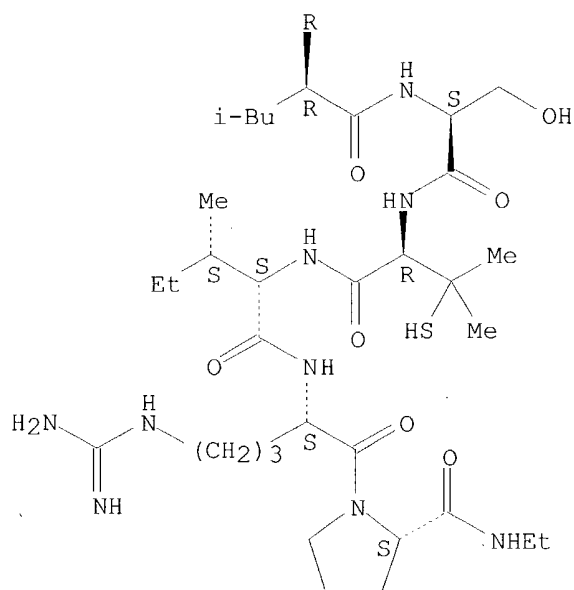
REFERENCE 1: 135:5822

L25 ANSWER 28 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251585-42-9 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-3-  
 mercapto-L-valyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H81 N13 O11 S  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

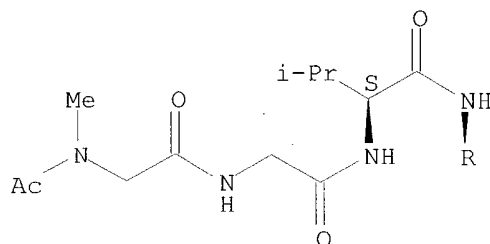
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

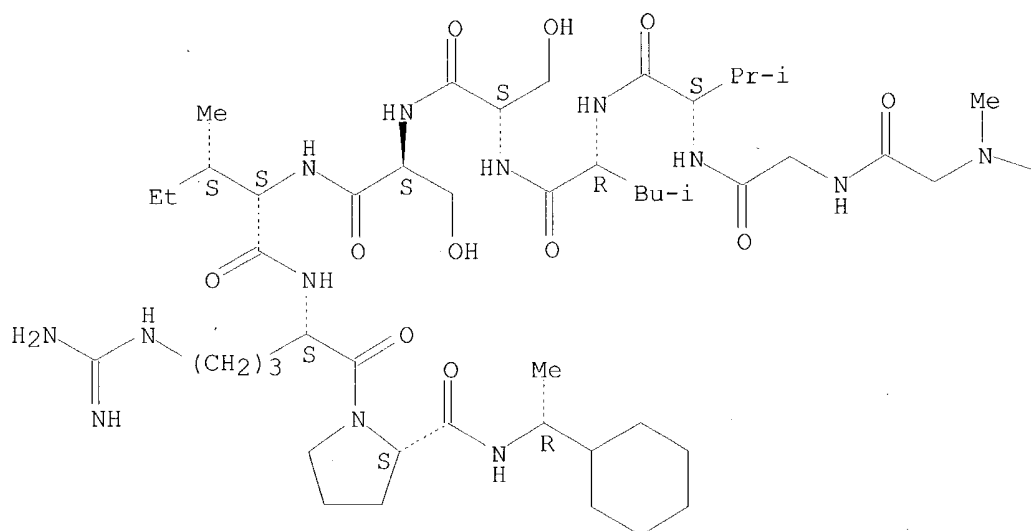
REFERENCE 1: 132:23191

L25 ANSWER 29 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-85-7 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-seryl-L-isoleucyl-L-arginyl-N-[(1R)-1-cyclohexylethyl]- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C49 H87 N13 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

Ac

1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

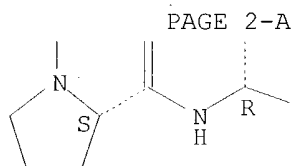
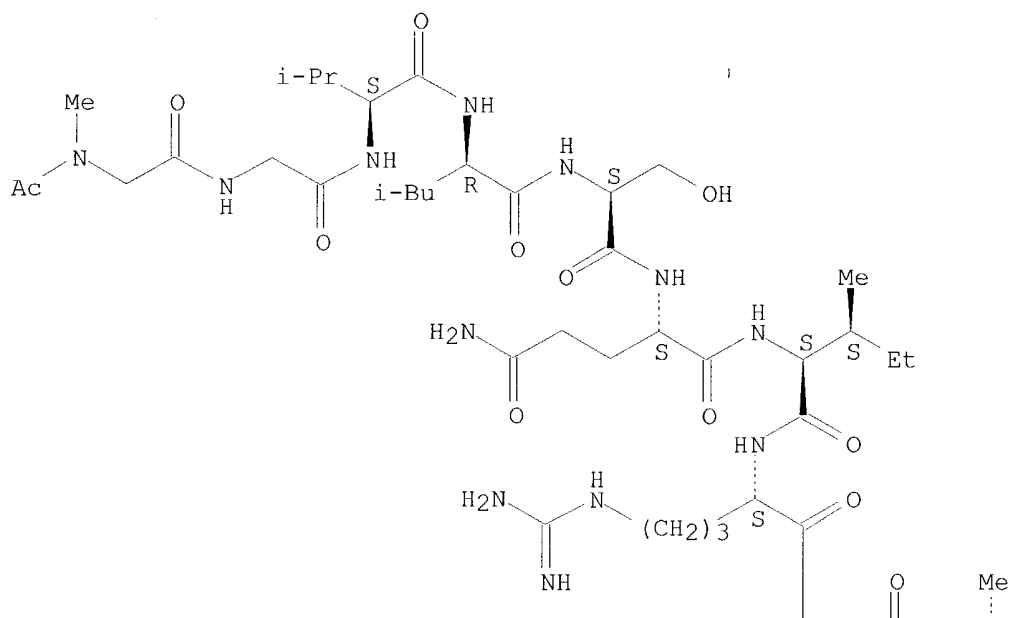
REFERENCE 1: 132:23191

L25 ANSWER 30 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-82-4 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-glutaminyl-L-isoleucyl-L-arginyl-N-[(1R)-1-cyclohexylethyl]- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H90 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

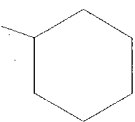
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.





PAGE 2-B

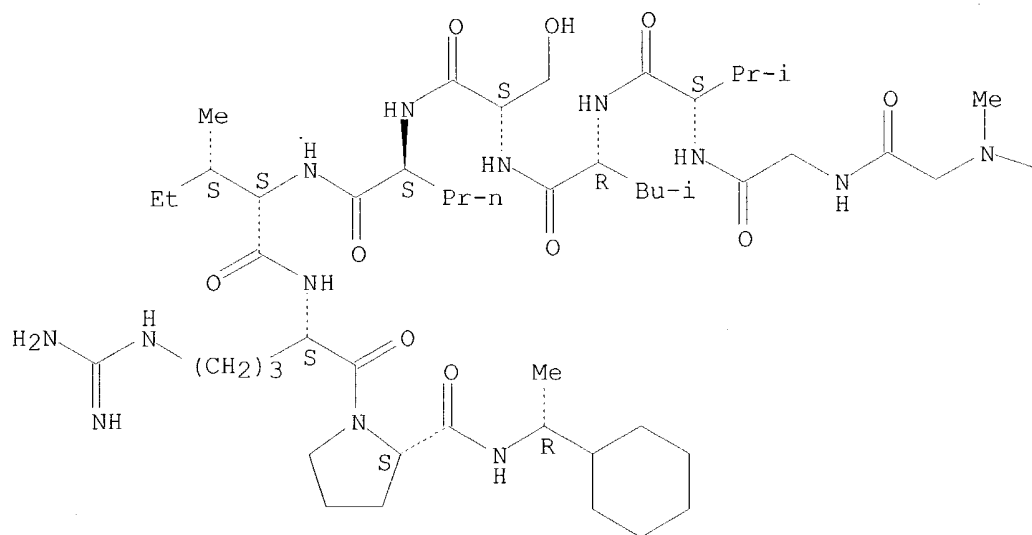


1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

```
L25 ANSWER 31 OF 67  REGISTRY  COPYRIGHT 2003 ACS on STN
RN 251584-81-3  REGISTRY
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-
norvalyl-L-isoleucyl-L-arginyl-N-[(1R)-1-cyclohexylethyl]- (9CI) (CA
INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C51 H91 N13 O11
SR CA
LC STN Files:  CA, CAPLUS, TOXCENTER
```

PAGE 1-A



PAGE 1-B

Ac

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

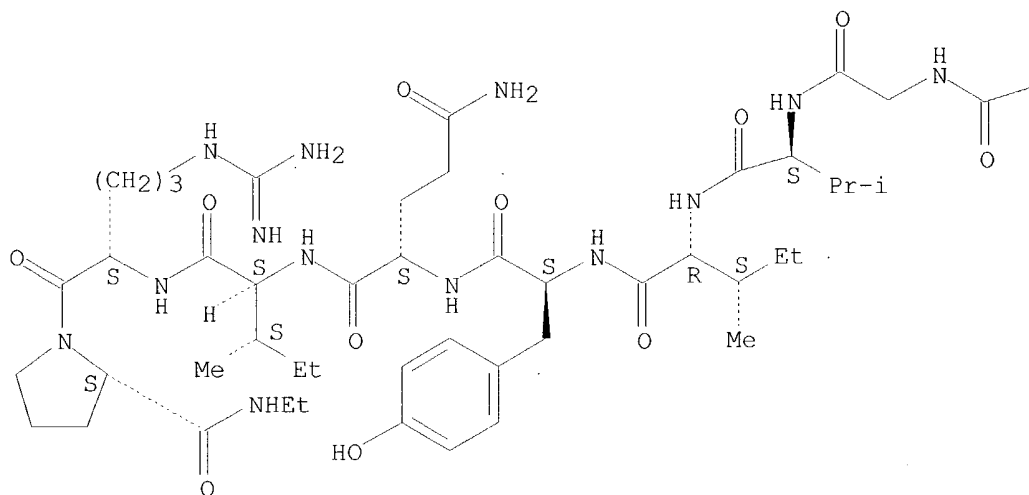
REFERENCE 1: 132:23191

```
L25 ANSWER 32 OF 67  REGISTRY  COPYRIGHT 2003 ACS on STN
RN 251584-55-1  REGISTRY
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-
tyrosyl-L-glutaminyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI)  (CA INDEX NAME)
FS PROTEIN SEQUENCE; STEREOSEARCH
MF C51 H84 N14 O12
SR CA
LC STN Files:  CA, CAPLUS, TOXCENTER
```

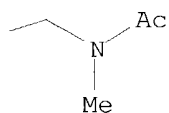
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Page 36

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

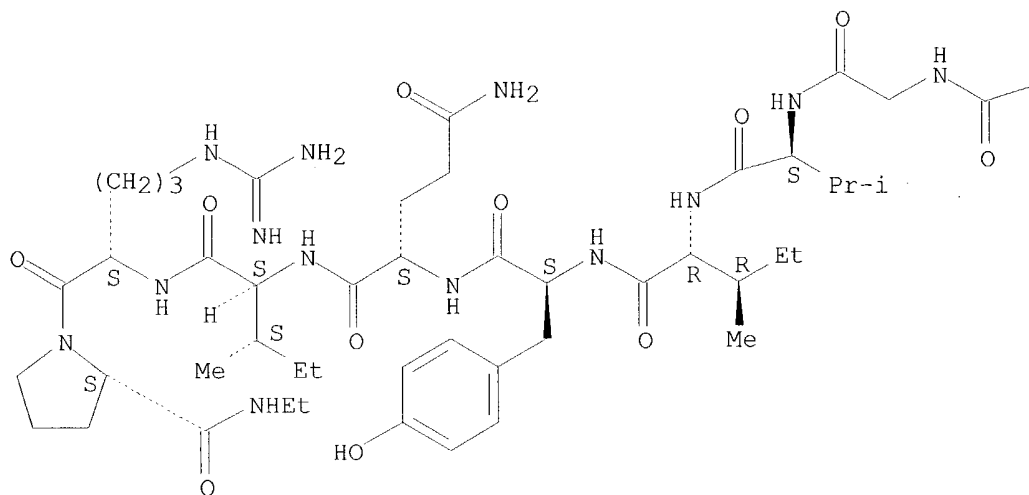
REFERENCE 1: 132:23191

L25 ANSWER 33 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-54-0 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-tyrosyl-  
 L-glutaminyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H84 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

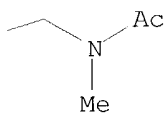
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



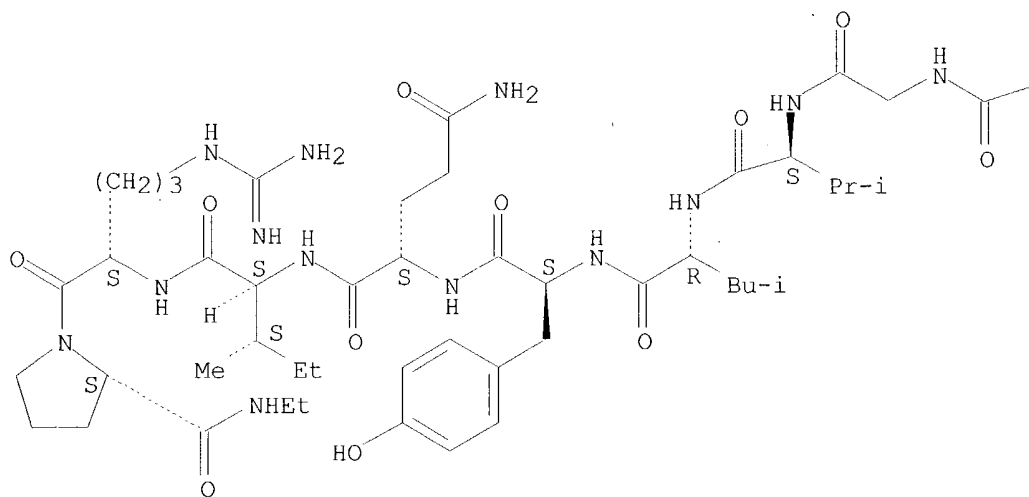
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

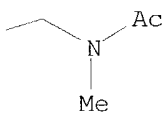
L25 ANSWER 34 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-53-9 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-tyrosyl-L-glutamyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H84 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

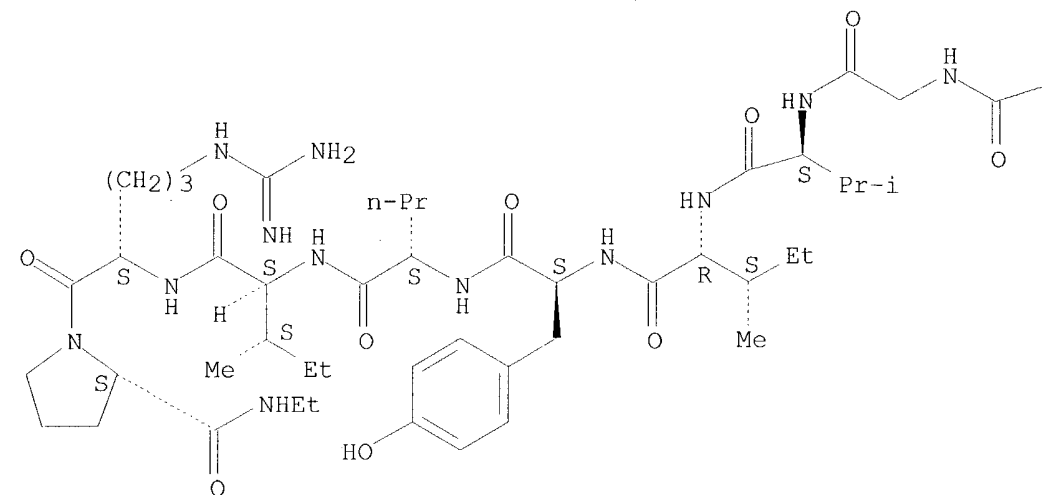
REFERENCE 1: 132:23191

L25 ANSWER 35 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-52-8 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-tyrosyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H85 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

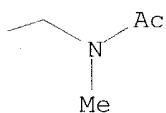
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

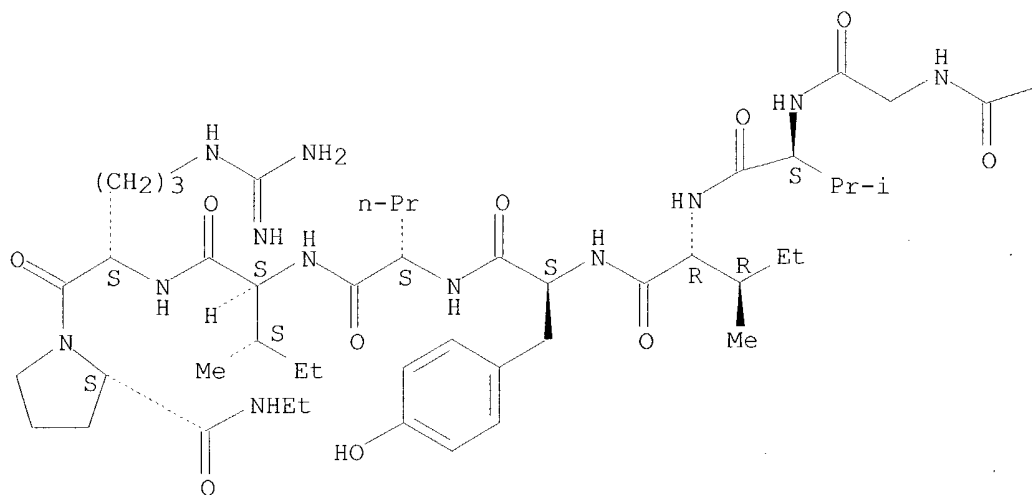
REFERENCE 1: 132:23191

L25 ANSWER 36 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-51-7 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-tyrosyl-  
 L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H85 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

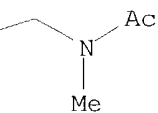
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



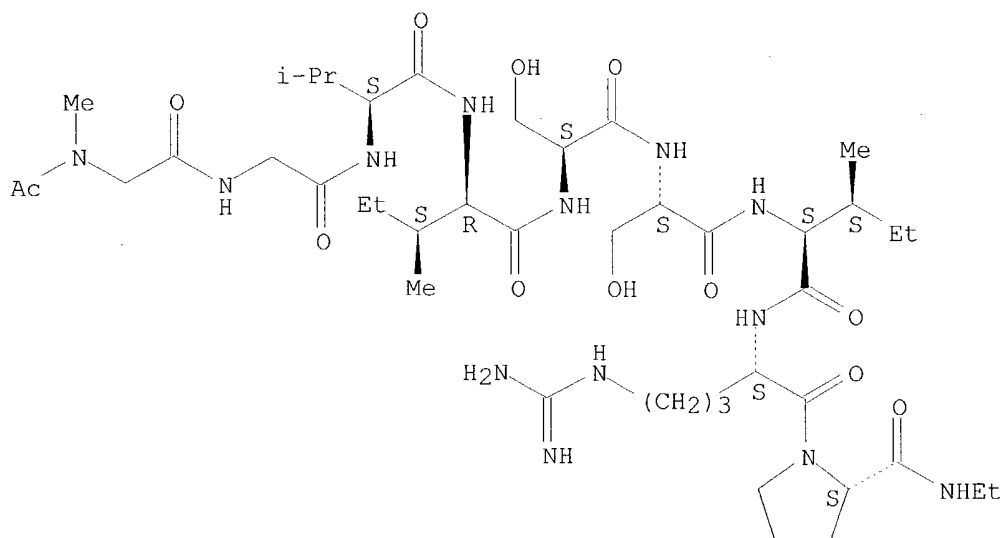
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 37 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-45-9 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-  
 seryl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C43 H77 N13 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 38 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 251584-44-8 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-L-leucyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

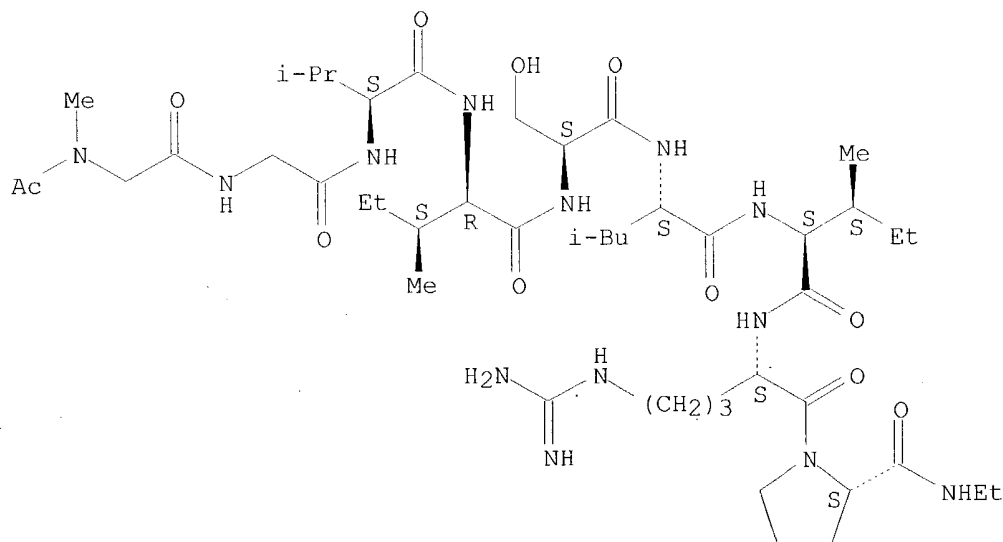
MF C46 H83 N13 O11

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.





1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

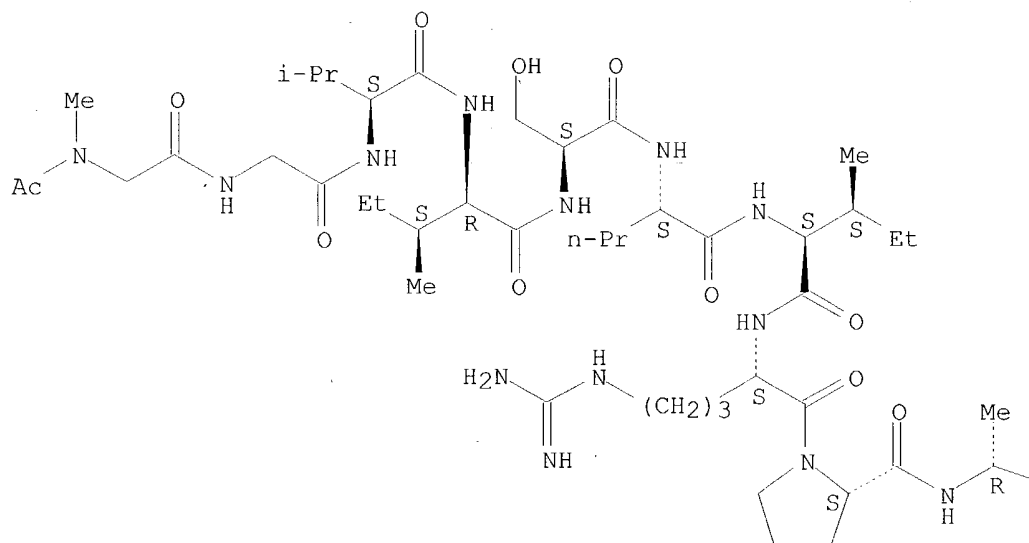
REFERENCE 1: 132:23191

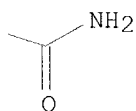
```
L25  ANSWER 39 OF 67  REGISTRY  COPYRIGHT 2003 ACS on STN
RN   251584-43-7  REGISTRY
CN   D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-
      seryl-L-norvalyl-L-isoleucyl-L-arginyl-L-prolyl- (9CI)  (CA INDEX NAME)
FS   PROTEIN SEQUENCE; STEREOSEARCH
MF   C46 H82 N14 O12
SR   CA
LC   STN Files:  CA, CAPLUS, TOXCENTER
```

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A





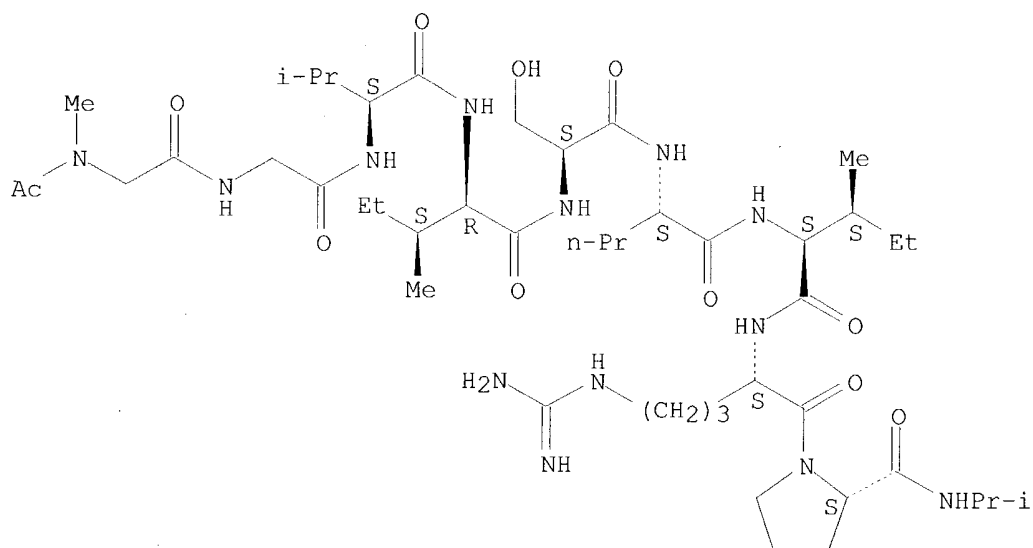
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 40 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-42-6 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-(1-methylethyl)- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C46 H83 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



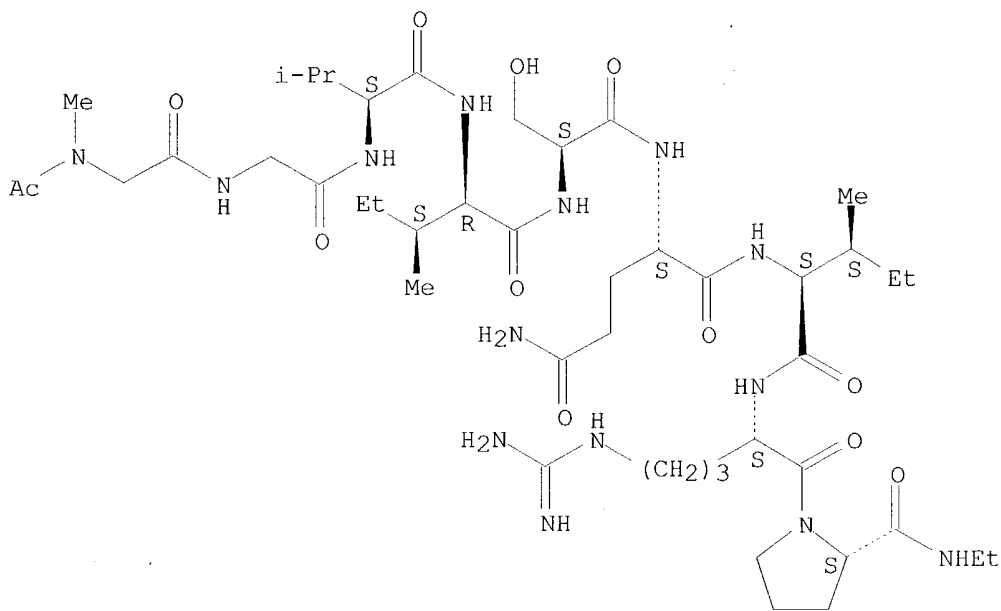
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 41 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-40-4 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-L-glutaminyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



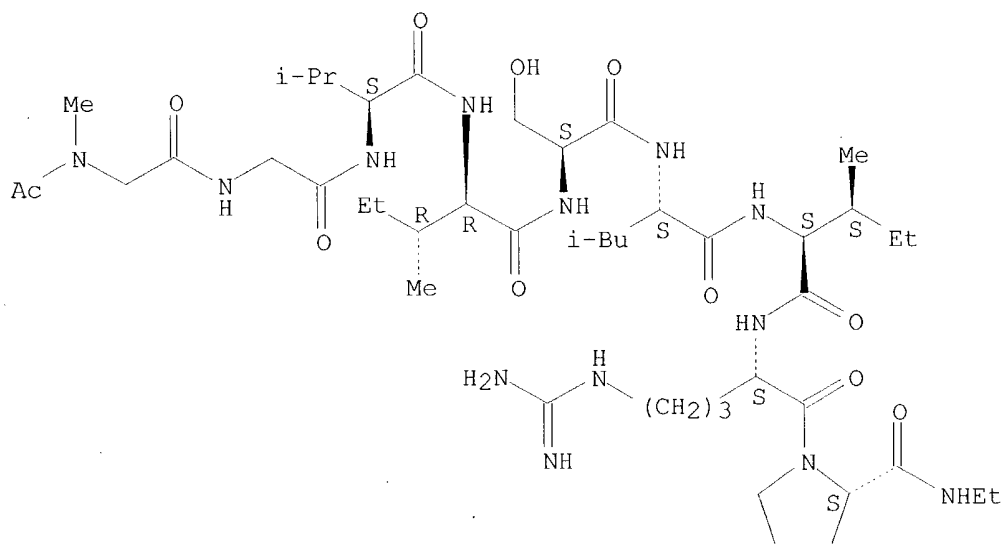
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 42 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-39-1 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-leucyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

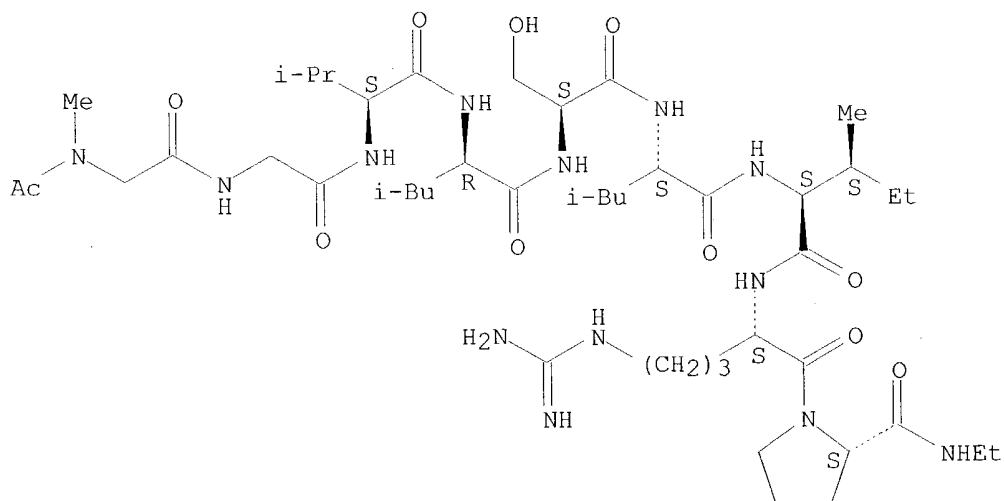


1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 43 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-38-0 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-leucyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C46 H83 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.



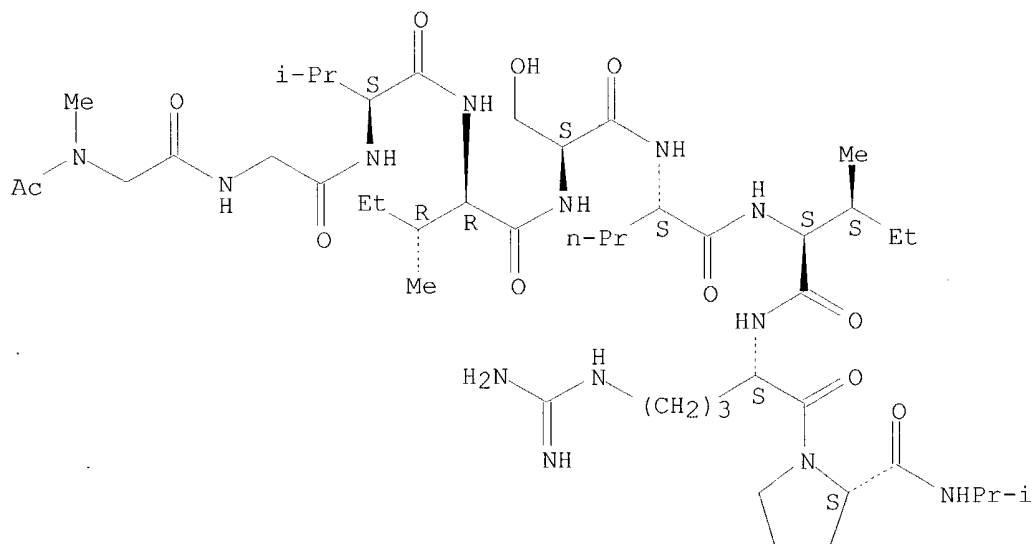
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 44 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-37-9 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-(1-methylethyl)- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



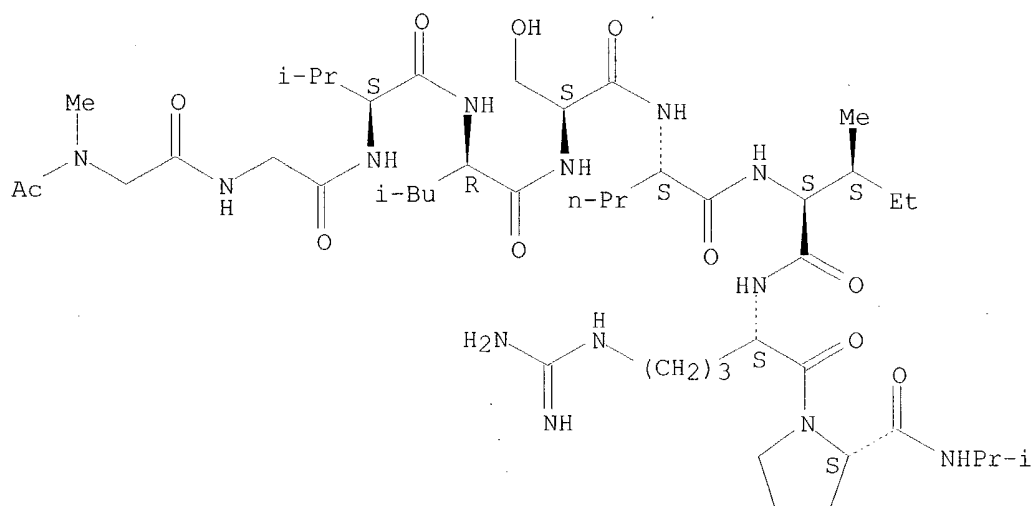
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 45 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-36-8 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-(1-methylethyl)- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

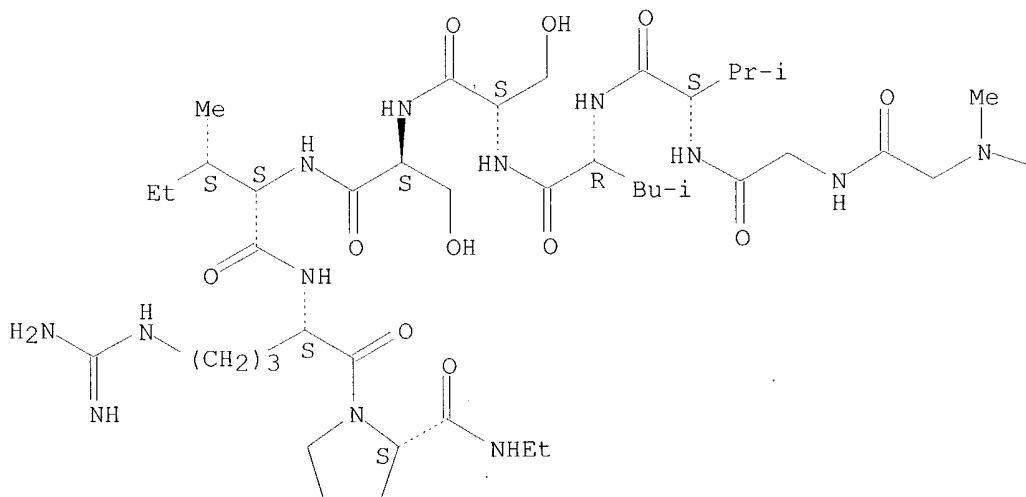
REFERENCE 1: 132:23191

L25 ANSWER 46 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-35-7 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C43 H77 N13 O12  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



Ac

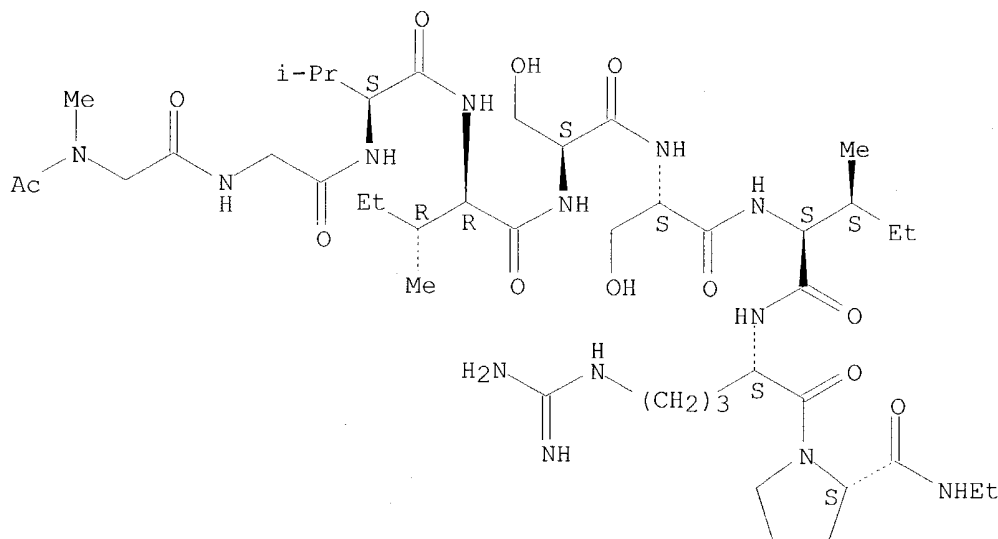
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 47 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-34-6 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-seryl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C43 H77 N13 O12  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

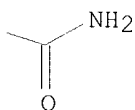
REFERENCE 1: 132:23191

L25 ANSWER 48 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-29-9 REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C46 H82 N14 O12  
SR CA



Absolute stereochemistry.

PAGE 1-B



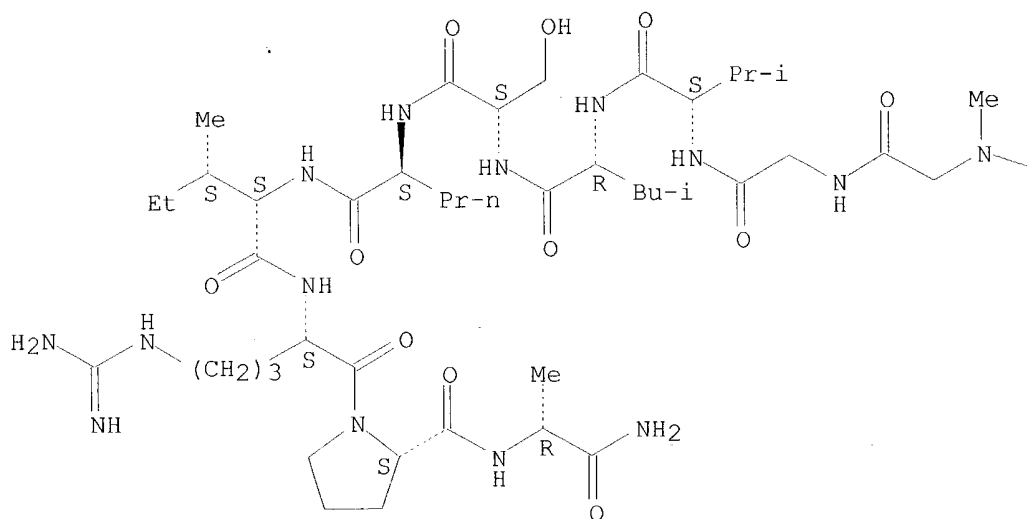
REFERENCE 1: 132:23191

Page 51

CN D-Alaninamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H82 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B

Ac

1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

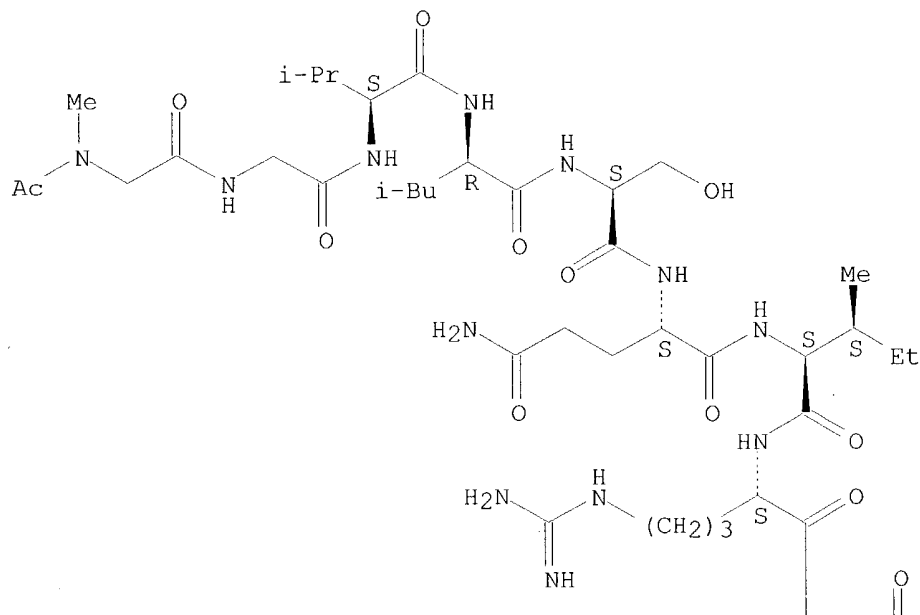
REFERENCE 1: 132:23191

L25 ANSWER 50 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-27-7 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-glutamyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

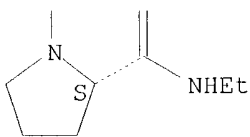
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



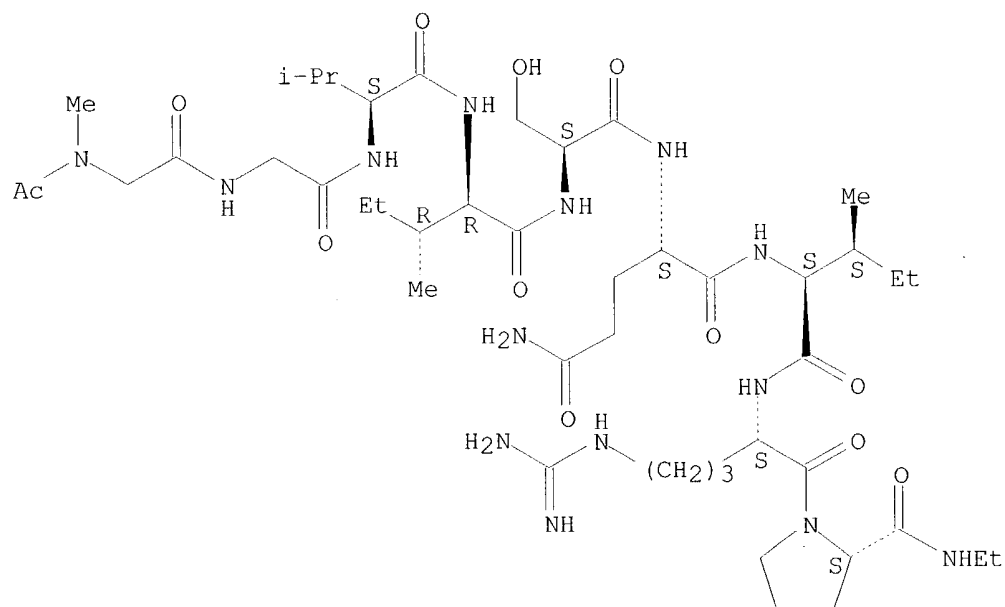
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 51 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-26-6 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-glutaminyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H80 N14 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



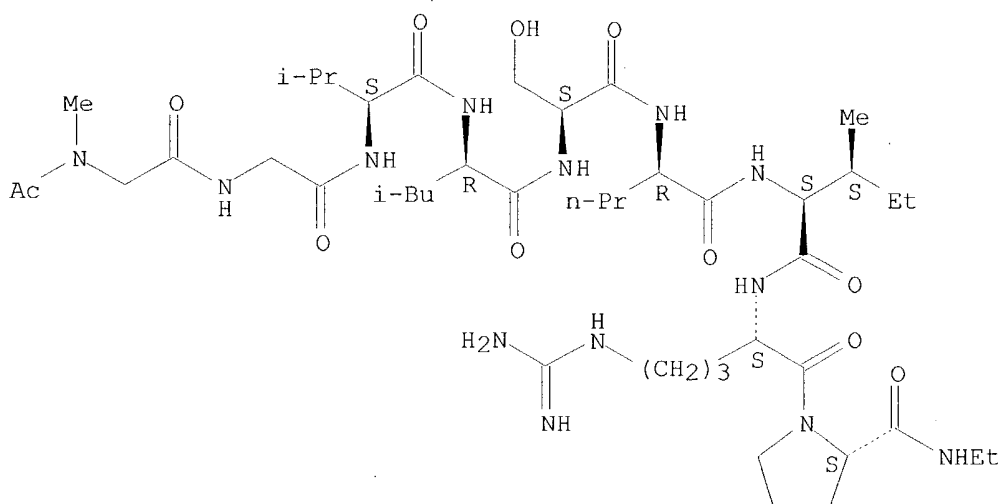
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 52 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-25-5 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-D-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H81 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



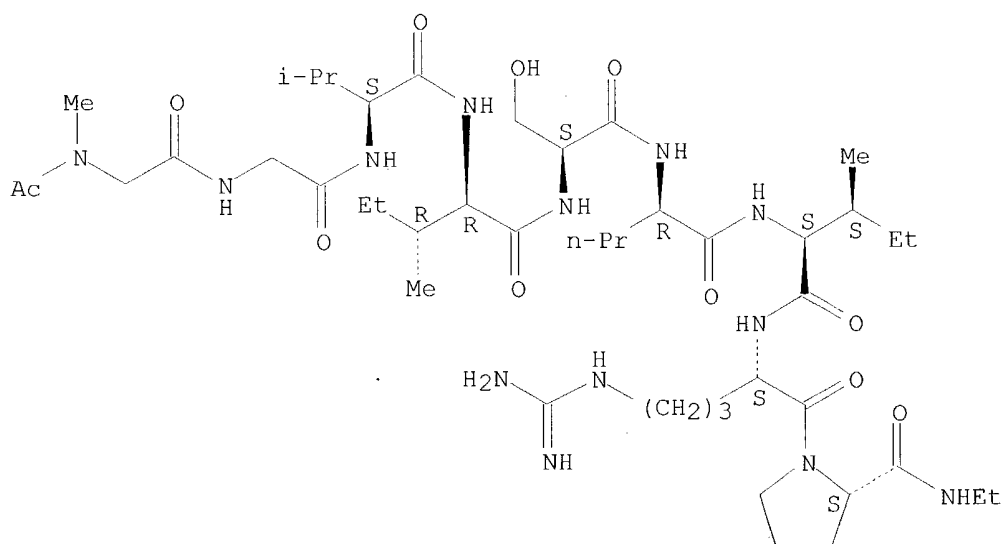
.1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 53 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-24-4 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-D-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H81 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

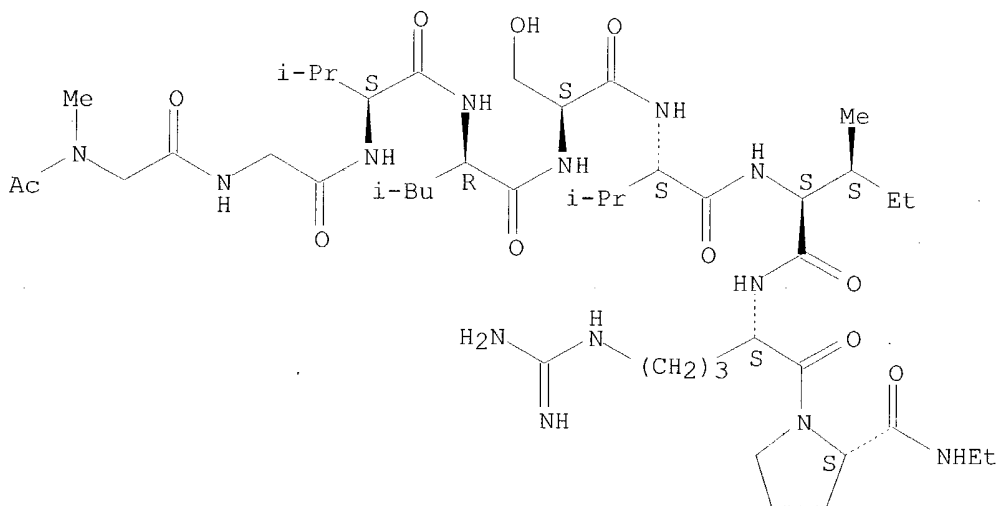


1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 54 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-18-6 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-valyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H81 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

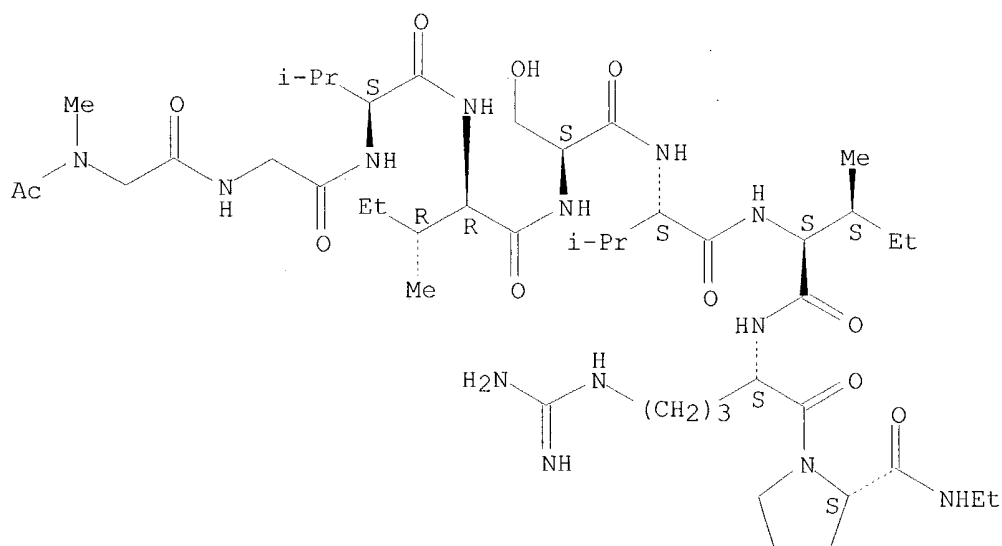


1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 55 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251584-17-5 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-valyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C45 H81 N13 O11  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 56 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 251584-11-9 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-alanyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

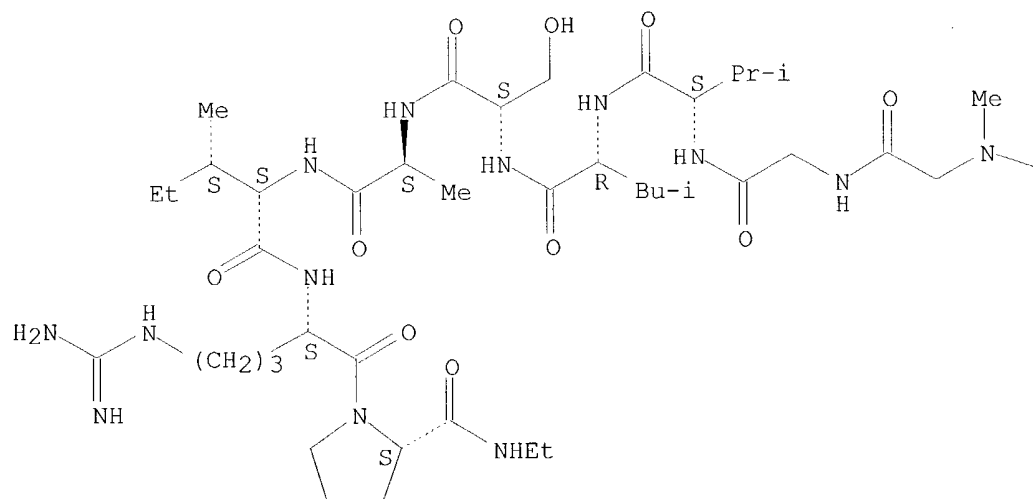
MF C43 H77 N13 O11

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.

PAGE 1-A



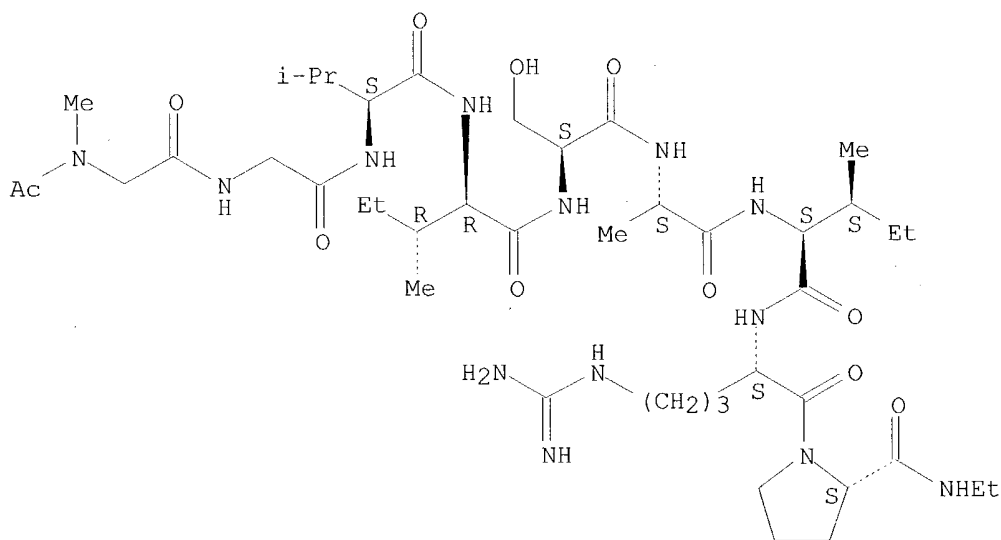
Ac

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 57 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251584-10-8 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-alanyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C43 H77 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

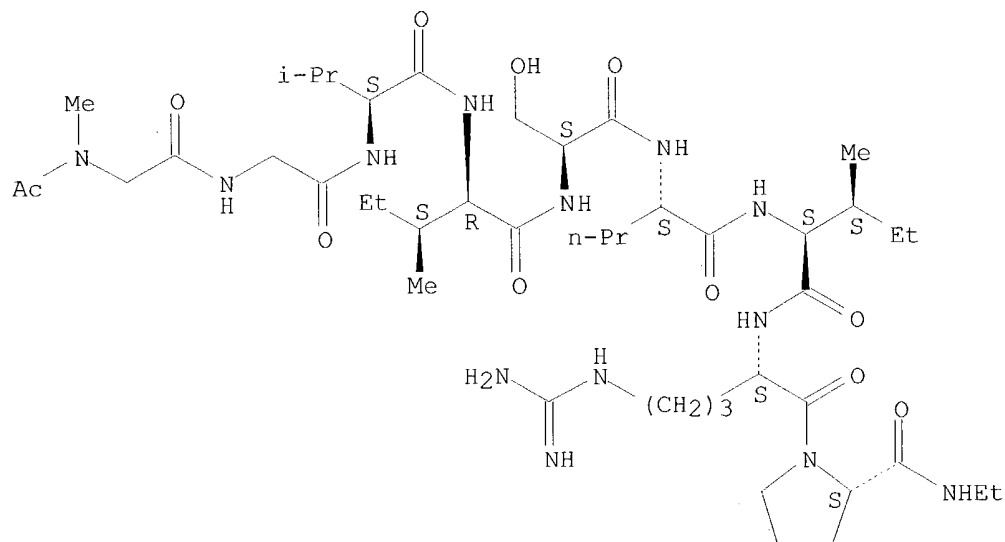
REFERENCE 1: 132:23191

L25 ANSWER 58 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251583-96-7 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-alloisoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H81 N13 O11  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER



\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



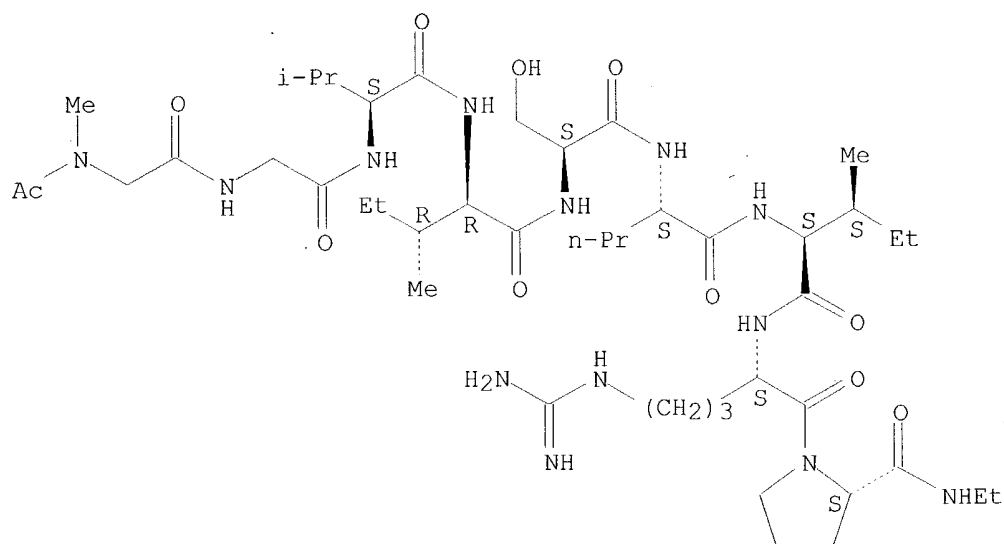
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 59 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251583-95-6 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H81 N13 O11  
SR CA

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

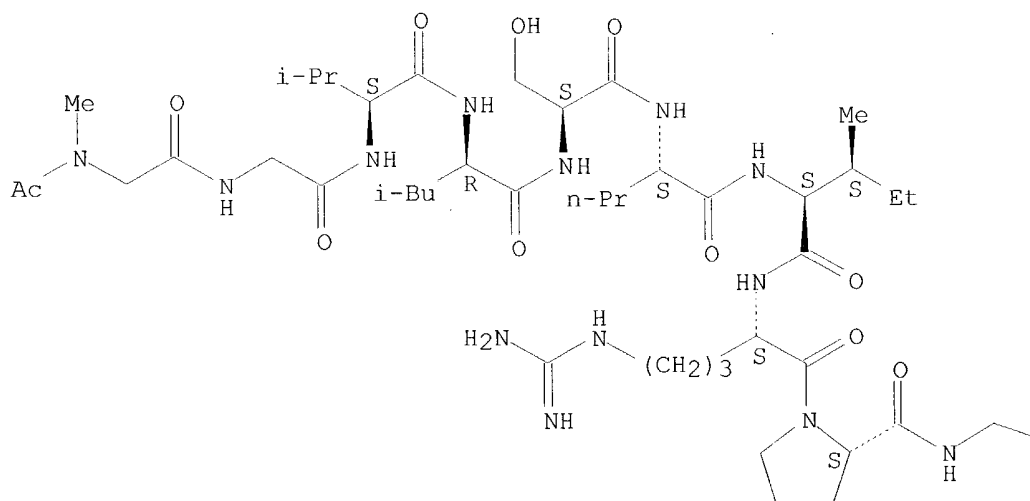


L25 ANSWER 60 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251583-86-5 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-(2-methoxyethyl)- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C46 H83 N13 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A





1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

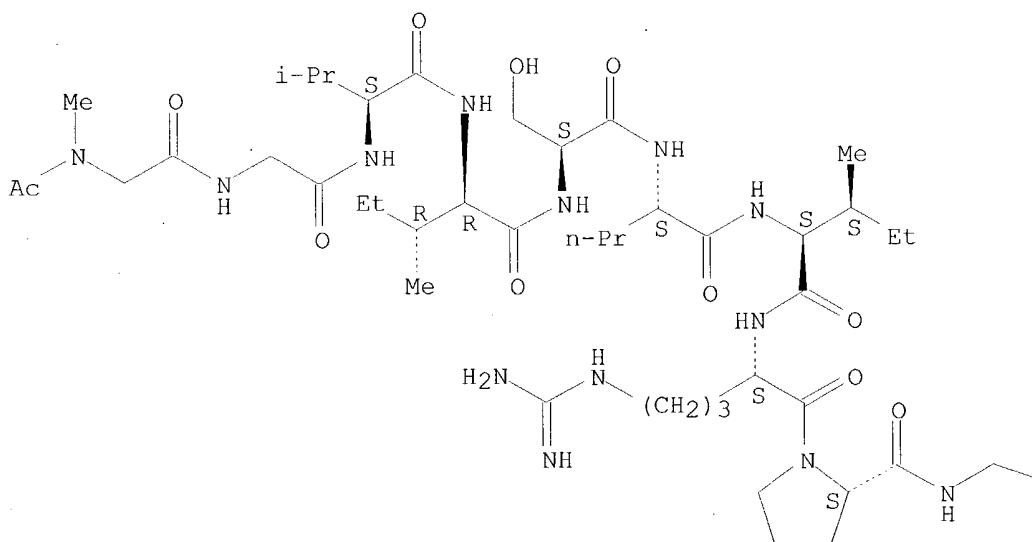
REFERENCE 1: 132:23191

L25 ANSWER 61 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251583-85-4 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-isoleucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-(2-methoxyethyl)- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C46 H83 N13 O12  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



OMe

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 62 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 251581-90-5 REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (salt)  
(9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C45 H81 N13 O11 . C2 H F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER

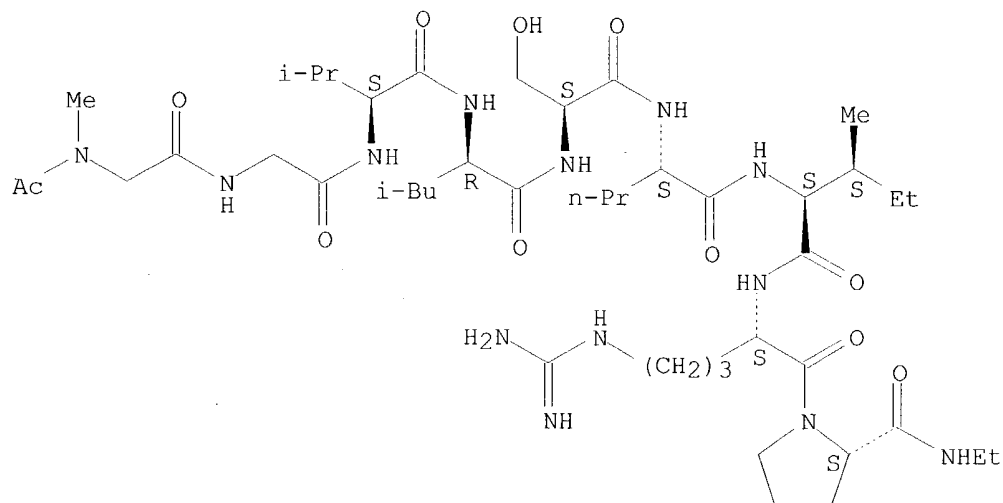
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 251581-89-2  
CMF C45 H81 N13 O11

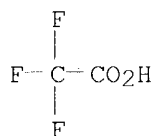
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



CM 2

CRN 76-05-1  
CMF C2 H F3 O2



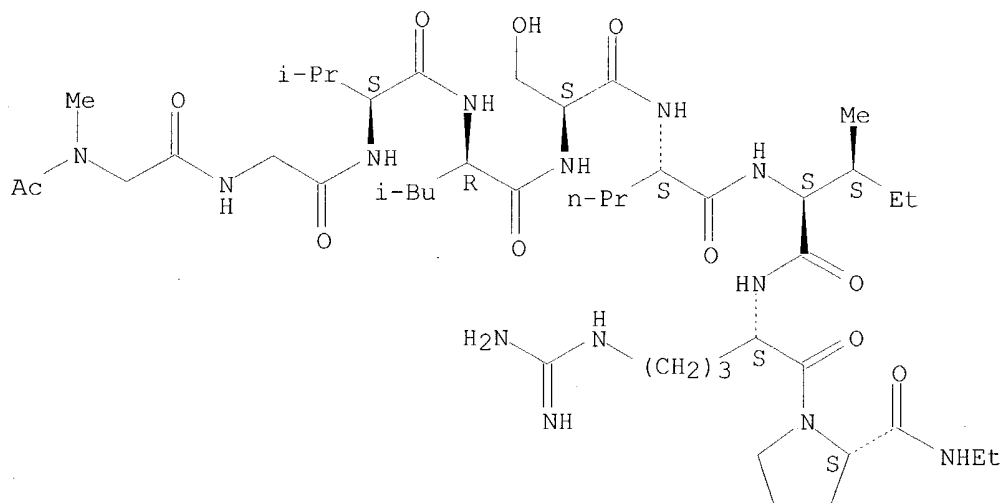
1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

```
L25  ANSWER 63 OF 67  REGISTRY  COPYRIGHT 2003 ACS on STN
RN   251581-89-2  REGISTRY
CN   L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-seryl-L-
     norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI)  (CA INDEX NAME)
FS   PROTEIN SEQUENCE; STEREOSEARCH
MF   C45 H81 N13 O11
CI   COM
SR   CA
LC   STN Files:    CA, CAPLUS, TOXCENTER
```

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 64 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251581-88-1 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-methionyl-  
 L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (9CI)  
 (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C47 H85 N13 O10 S . C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

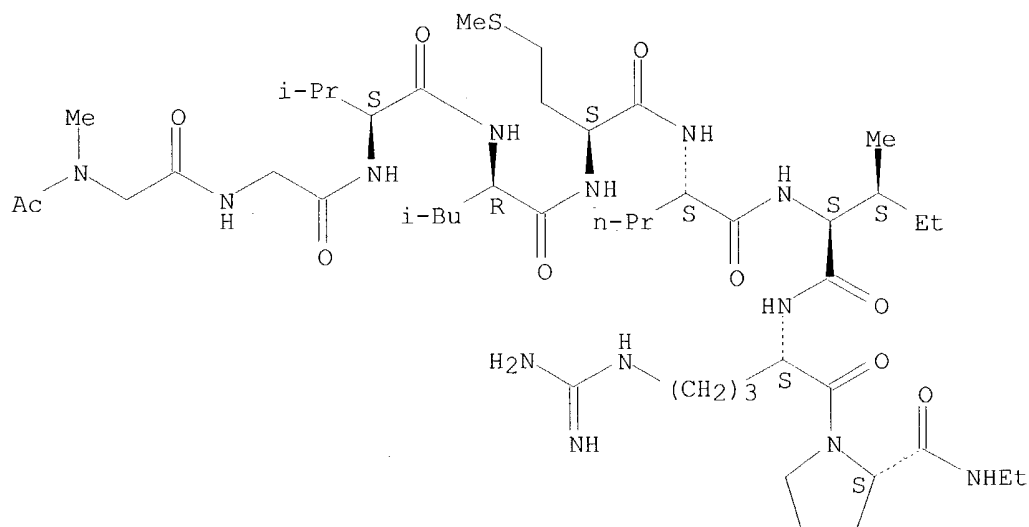
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 251581-87-0  
 CMF C47 H85 N13 O10 S

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

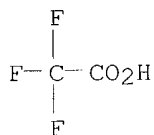
Absolute stereochemistry.



CM 2

CRN 76-05-1

CMF C2 H F3 O2



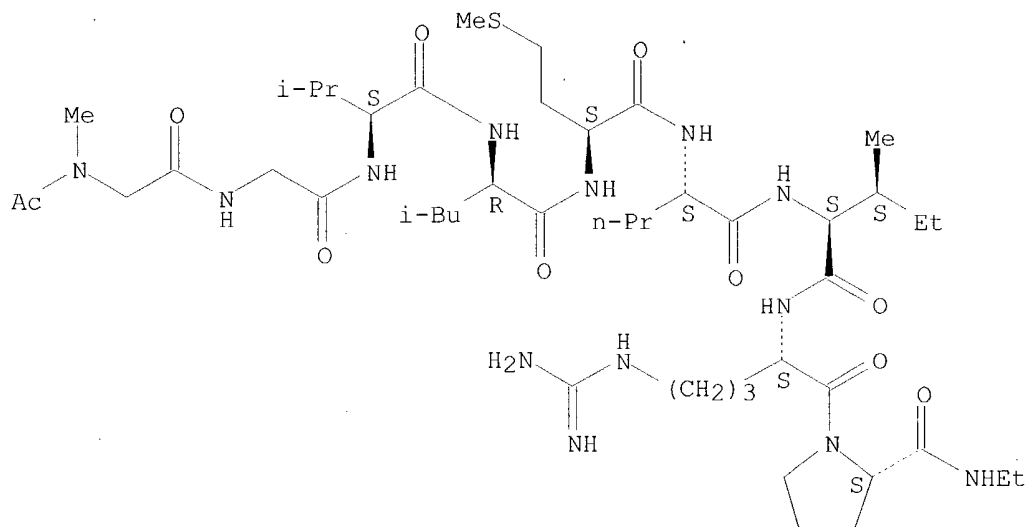
1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 65 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251581-87-0 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-methionyl-  
 L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C47 H85 N13 O10 S  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

L25 ANSWER 66 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN

RN 251581-72-3 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-tyrosyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C51 H85 N13 O11 . C2 H F3 O2

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

CRN 251581-71-2

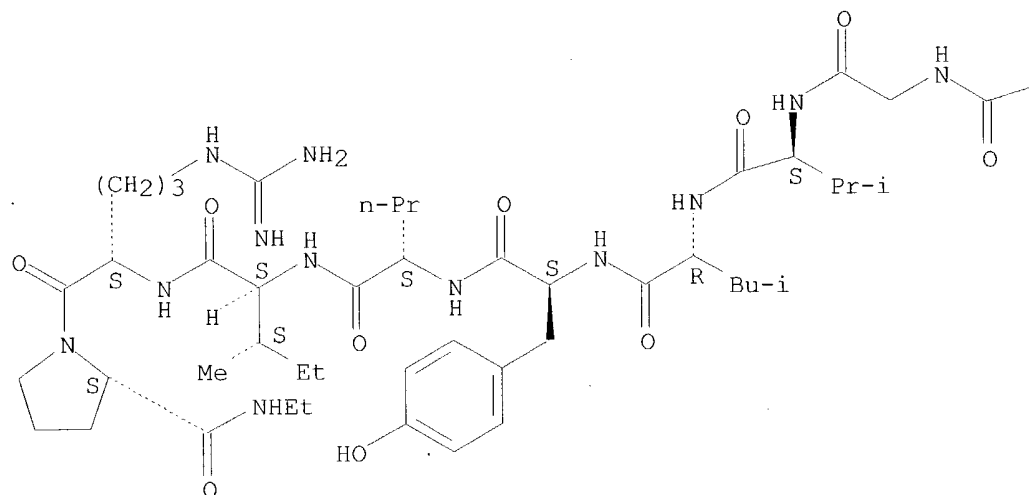
CMF C51 H85 N13 O11

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

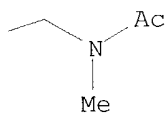
Absolute stereochemistry.



PAGE 1-A



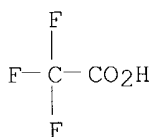
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

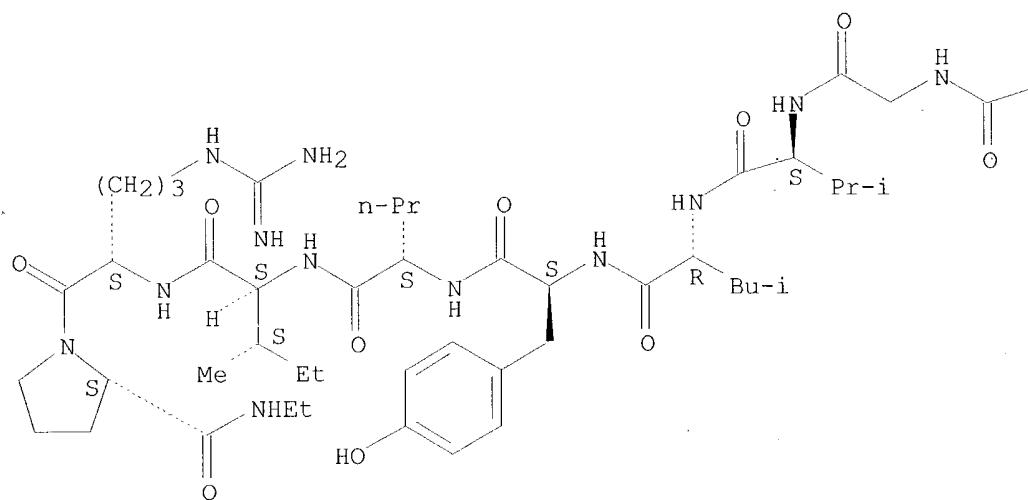
REFERENCE 1: 132:23191

L25 ANSWER 67 OF 67 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN 251581-71-2 REGISTRY  
 CN L-Prolinamide, N-acetyl-N-methylglycylglycyl-L-valyl-D-leucyl-L-tyrosyl-L-norvalyl-L-isoleucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C51 H85 N13 O11  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, TOXCENTER

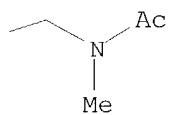
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:23191

=&gt; fil hcaplus

FILE 'HCAPLUS' ENTERED AT 16:25:28 ON 28 SEP 2003  
 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
 PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
 COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

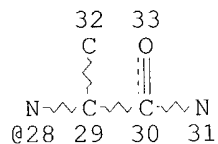
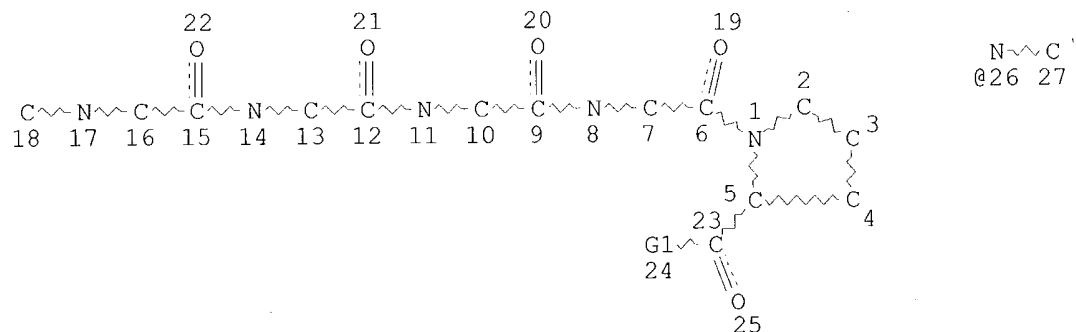
FILE COVERS 1907 - 28 Sep 2003 VOL 139 ISS 14  
 FILE LAST UPDATED: 26 Sep 2003 (20030926/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=>  
 =>

=&gt; d stat que 130

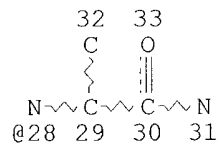
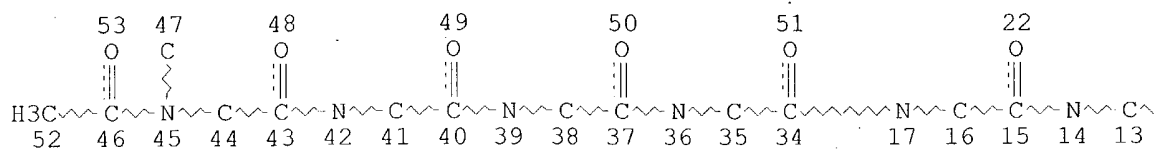
L1 STR



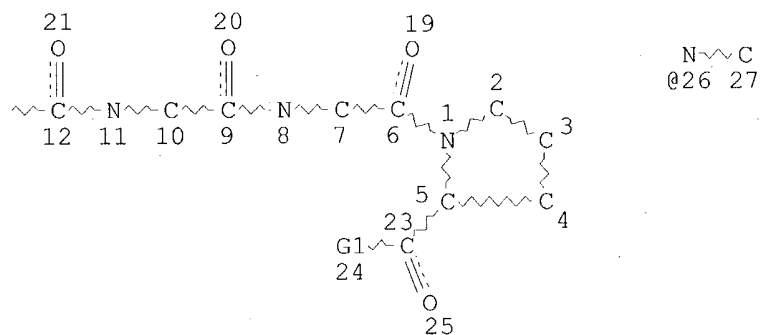
VAR G1=26/28  
 NODE ATTRIBUTES:  
 DEFAULT MLEVEL IS ATOM  
 DEFAULT ECLEVEL IS LIMITED

GRAPH ATTRIBUTES:  
 RING(S) ARE ISOLATED OR EMBEDDED  
 NUMBER OF NODES IS 33

STEREO ATTRIBUTES: NONE  
 L7 121997 SEA FILE=REGISTRY SSS FUL L1  
 L22 STR



Page 1-A



Page 1-B

VAR G1=26/28

NODE ATTRIBUTES:

DEFAULT MLEVEL IS ATOM

DEFAULT ECLEVEL IS LIMITED

GRAPH ATTRIBUTES:

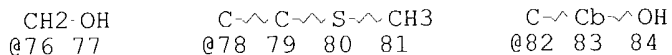
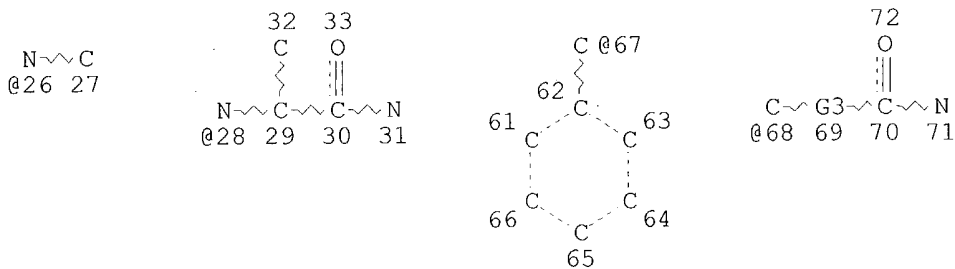
RING(S) ARE ISOLATED OR EMBEDDED

NUMBER OF NODES IS 52

STEREO ATTRIBUTES: NONE

L23 780 SEA FILE=REGISTRY SUB=L7 SSS FUL L22

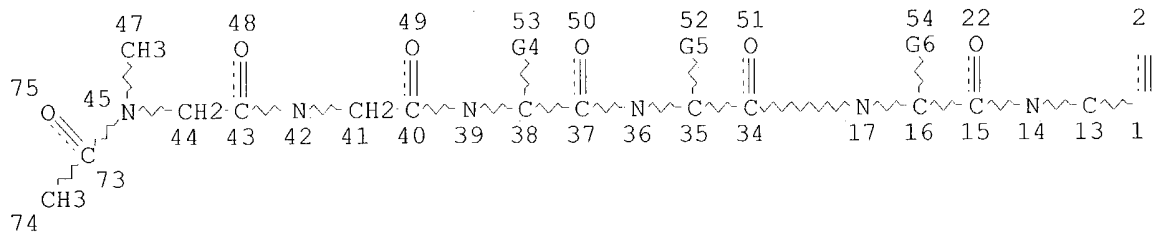
L24 STR



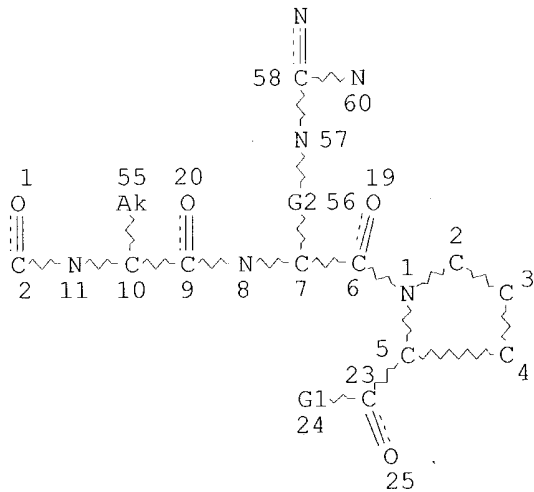
Page 1-A

59

Page 1-B



Page 2-A



Page 2-B

VAR G1=26/28

REP G2=(3-3) C

REP G3=(0-1) C

VAR G4=67/68/I-PR

VAR G5=ME/ET/I-PR/N-PR/I-BU/N-BU/T-BU/S-BU

VAR G6=76/78/82

NODE ATTRIBUTES:

DEFAULT MLEVEL IS ATOM

DEFAULT ECLEVEL IS LIMITED

GRAPH ATTRIBUTES:  
RING(S) ARE ISOLATED OR EMBEDDED  
NUMBER OF NODES IS 82

STEREO ATTRIBUTES: NONE

L25 67 SEA FILE=REGISTRY SUB=L7 SSS FUL L24  
L26 3 SEA FILE=HCAPLUS ABB=ON PLU=ON L25  
L27 713 SEA FILE=REGISTRY ABB=ON PLU=ON L23 NOT L25  
L28 23 SEA FILE=HCAPLUS ABB=ON PLU=ON L27  
L29 20 SEA FILE=HCAPLUS ABB=ON PLU=ON L28 NOT L26  
L30 18 SEA FILE=HCAPLUS ABB=ON PLU=ON L29 AND PD<=APRIL 11, 2001

=>

=>

=> d ibib abs hitrn 130 1-18

L30 ANSWER 1 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2000:307128 HCAPLUS

DOCUMENT NUMBER: 132:322148

TITLE: Preparation of thrombin inhibitors based on the amino acid sequence of hirudin

INVENTOR(S): Dimaio, John; Konishi, Yasuo; Ni, Feng; Steinmetzer, Torsten

PATENT ASSIGNEE(S): The National Research Council of Canada, Can.

SOURCE: U.S., 49 pp., Cont.-in-part of U.S. Ser. No. 302,245, abandoned.

CODEN: USXXAM

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6060451	A	20000509	US 1995-406142	19950320 <--
CA 2215702	AA	19960926	CA 1996-2215702	19960318 <--
WO 9629347	A1	19960926	WO 1996-CA164	19960318 <--
W: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI				
RW: KE, LS, MW, SD, SZ, UG, AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN				
AU 9649349	A1	19961008	AU 1996-49349	19960318 <--
AU 695920	B2	19980827		
EP 815139	A1	19980107	EP 1996-905636	19960318 <--
EP 815139	B1	20011107		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
CN 1182436	A	19980520	CN 1996-193457	19960318 <--
BR 9607839	A	19980616	BR 1996-7839	19960318 <--
JP 11502203	T2	19990223	JP 1996-527932	19960318 <--
IL 117526	A1	19991231	IL 1996-117526	19960318 <--
AT 208401	E	20011115	AT 1996-905636	19960318
ES 2168461	T3	20020616	ES 1996-905636	19960318
ZA 9602267	A	19960927	ZA 1996-2267	19960320 <--
NO 9704342	A	19971119	NO 1997-4342	19970919 <--
HK 1005511	A1	20020315	HK 1998-104773	19980603
PRIORITY APPLN. INFO.:			US 1994-302245	B2 19940908
			US 1995-406142	A 19950320

WO 1996-CA164 W 19960318

OTHER SOURCE(S): MARPAT 132:322148

AB Thrombin inhibitors AS-Y-Z-A [AS is a hydrophobic moiety which binds the catalytic site of thrombin and which comprises (a) one or two hydrophobic .alpha.-amino acids which are optionally substituted by alkyl, aryl, or aralkyl and (b) a guanidino group; Y = CO, CH<sub>2</sub>, CH<sub>2</sub>OH; Z is a divalent, straight-chained linker moiety that has a chain length of approx. 10-85 atoms; A is an acidic portion of formula -G-X'-G'-Q-Q1-Q2(W')-, where G and G' are each an L-.alpha.-amino acid having pk value .ltoreq. 5, X' is a hydrophobic L-.alpha.-amino acid, Q is and L-.alpha.-amino acid or a cyclic L-imino acid; Q1 and Q2 are different and are either Ile or Pro; W' is H, alkyl, aryl, or aralkyl, with the proviso that W' is linked to whichever of Q1 or Q2 is Pro] and its pharmaceutically acceptable salts were prepd. for treatment of thrombotic disorders. Thus, Ac-D-Phe-Pro-Arg-.PSI.[COCH<sub>2</sub>]CH<sub>2</sub>CO-Gln-Ser-His-Asn-Asp-Gly-Asp-Phe-Glu-Glu-Ile-Pro-Glu-Glu-Tyr-Leu-Gln-OH (P79) was prepd. by the solid phase method and tested for thrombin inhibitory activity (IC<sub>50</sub> = 2 nM in the platelet aggregation test).

IT 183969-25-7P, P536

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(prepn. of thrombin inhibitors based on the amino acid sequence of hirudin)

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L30 ANSWER 2 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2000:144132 HCAPLUS

DOCUMENT NUMBER: 132:152142

TITLE: Synthesis of peptides with N-substituted glycines as luteinizing hormone-releasing hormone inhibitory analogs for treatment of hormone-dependent tumors.

INVENTOR(S): Dechantsreiter, Michael; Kessler, Horst; Bernd, Michael; Kutscher, Bernhard; Beckers, Thomas

PATENT ASSIGNEE(S): Asta Medica A.-G., Germany

SOURCE: Ger. Offen., 32 pp.

CODEN: GWXXBX

DOCUMENT TYPE: Patent

LANGUAGE: German

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
DE 19941248	A1	20000302	DE 1999-19941248	19990831 <--
			DE 1998-19839817	19980901

PRIORITY APPLN. INFO.:

OTHER SOURCE(S): MARPAT 132:152142

AB Title decapeptide compds. in which one or two glycine amine groups have been substituted with side-chain equiv. of natural or non-natural amino acids were prepd. as analogs of LH-RH, for use in treating hormone-dependent tumors or for LH-RH suppression therapies (no data). Thus, amino acid substitutes were prepd. by, for example, alkylation of an amine such as 4-Cl-C<sub>6</sub>H<sub>4</sub>-NH<sub>2</sub> with BrCH<sub>2</sub>COOEt, or amination of CHOCO<sub>2</sub>H with RNH(CH<sub>2</sub>)<sub>2</sub>OC(CH<sub>3</sub>)<sub>3</sub> (R = protecting group). The amino acid substitutes could then be used in solid-phase synthesis (BOC or Fmoc chem.) to prep. fragments for soln. coupling to give the final decapeptides.

IT 258332-71-7P

RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)

(synthesis of N-substituted glycines for use in prepn. of peptides as LH-releasing hormone inhibitory analogs for treatment of hormone-dependent tumors)

IT 258332-86-4P

RL: SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(synthesis of N-substituted glycines for use in prepn. of peptides as LH-releasing hormone inhibitory analogs for treatment of hormone-dependent tumors)

L30 ANSWER 3 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1999:439184 HCAPLUS

DOCUMENT NUMBER: 131:237478

TITLE: Potent Bivalent Thrombin Inhibitors: Replacement of the Scissile Peptide Bond at P1-P1' with Arginyl Ketomethylene Isosteres

AUTHOR(S): Steinmetzer, Torsten; Zhu, Bing Yan; Konishi, Yasuo

CORPORATE SOURCE: Biotechnology Research Institute, National Research Council Canada, Montreal, QC, H4P 2R2, Can.

SOURCE: Journal of Medicinal Chemistry (1999), 42(16), 3109-3115

CODEN: JMCMAR; ISSN: 0022-2623

PUBLISHER: American Chemical Society

DOCUMENT TYPE: Journal

LANGUAGE: English

OTHER SOURCE(S): CASREACT 131:237478

AB The authors have designed highly potent synthetic bivalent thrombin inhibitors, which consist of an active site blocking segment, a fibrinogen recognition exosite blocking segment, and a linker connecting these segments. The bivalent inhibitors bind to the active site and the fibrinogen recognition exosite simultaneously. As a result, the inhibitors showed much higher affinity for thrombin than the individual blocking segments. Various arginyl ketomethylene isosteres Arg.PSI.[CO-CH2-X]P1' were incorporated into the bivalent inhibitors as P1-P1' segment to eliminate the scissile bond. The P1' residue is a natural or unnatural amino acid; specifically, the incorporation of mercaptoacetic acid exhibited superiority in synthesis and affinity for thrombin. Inhibitor (D-cyclohexylalanine)-Pro-Arg.PSI.[CO-CH2-S]Gly-(Gly)4-Asp-Tyr-Glu-Pro-Ile-Pro-Glu-Glu-Tyr-cyclohexylalanine-(D-Glu)-OH showed the lowest Ki value of 3.5.times.10<sup>-13</sup> M, which is comparable to that (Ki = 2.3.times.10<sup>-13</sup> M) of recombinant hirudin. Consequently the authors successfully reduced the size of the inhibitor from .apprx.7 kDa of recombinant hirudin to .apprx.2 kDa without losing the affinity.

IT 183969-25-7P

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(potent bivalent thrombin inhibitors in relation to replacement of scissile peptide bond at P1-P1' with arginyl ketomethylene isosteres)

REFERENCE COUNT: 51 THERE ARE 51 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L30 ANSWER 4 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1997:168540 HCAPLUS

DOCUMENT NUMBER: 126:152828

TITLE: LHRH antagonist synthetic peptide analogs for use as cancer inhibitors, contraceptives, or other pharmaceuticals

INVENTOR(S): Roeske, Roger W.

PATENT ASSIGNEE(S): Indiana University Foundation, USA; Roeske, Roger W.

SOURCE: PCT Int. Appl., 52 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:



PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9640757	A2	19961219	WO 1996-US9852	19960607 <--
WO 9640757	A3	19970220		
W: AU, CA, JP, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 5843901	A	19981201	US 1995-480494	19950607 <--
CA 2219460	AA	19961219	CA 1996-2219460	19960607 <--
AU 9661680	A1	19961230	AU 1996-61680	19960607 <--
AU 715399	B2	20000203		
EP 794961	A2	19970917	EP 1996-919311	19960607 <--
EP 794961	B1	20020828		
R: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LI, LU, MC, NL, PT, SE				
JP 11507374	T2	19990629	JP 1996-502050	19960607 <--
EP 1188768	A2	20020320	EP 2001-204149	19960607
EP 1188768	A3	20021023		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
AT 222924	E	20020915	AT 1996-919311	19960607
ES 2177789	T3	20021216	ES 1996-919311	19960607
US 6423686	B1	20020723	US 1998-973378	19980406
US 2002115615	A1	20020822		
US 6455499	B1	20020924	US 1999-256599	19990223
US 2003040482	A1	20030227	US 2002-115553	20020402
US 2003181385	A1	20030925	US 2002-117364	20020405
PRIORITY APPLN. INFO.:			US 1995-480494	A 19950607
			EP 1996-919311	A3 19960607
			WO 1996-US9852	W 19960607
			US 1998-973378	A3 19980406

OTHER SOURCE(S): MARPAT 126:152828

AB Many novel LH-releasing hormone(LHRH) antagonist peptide analogs or peptide mimetics, pharmaceutical compns. thereof, and methods of use thereof, are disclosed. The LHRH antagonist comprises a peptide compd., wherein a residue of the peptide compd. corresponding to the amino acid at position 6 of natural mammalian LHRH comprises a hydrophilic N-acyl moiety, a dipolar moiety, a sulfonium moiety, a receptor-modifying moiety or a small polar moiety. LHRH antagonist peptides are useful as inhibitors of sex hormone-dependent cancers (e.g., prostate cancer). LHRH antagonist peptides are also useful as contraceptive agents. The peptides can be used to treat other LHRH-related disorders as well, such as precocious puberty or premenstrual syndrome. The anti-ovulatory and histamine release activity of LHRH antagonists are compared. S.c. injections of LHRH antagonists suppressed plasma testosterone levels.

IT 186836-17-9P 186837-16-1P 186837-18-3P  
186837-28-5P 186837-31-0P

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(LHRH antagonist synthetic peptide analogs with pharmaceutical applications as cancer inhibitors or contraceptive agents)

L30 ANSWER 5 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1996:696046 HCAPLUS

DOCUMENT NUMBER: 126:42251

TITLE: LHRH antagonists conformationally restricted at the N-terminus

AUTHOR(S): Swenson, R. E.; Mort, N. A.; Haviv, F.; Nichols, C. J.; Bush, E. N.; Diaz, G. J.; Bammert, G. F.; Rhutasel, N. S.; Nguyen, A.; et al.

CORPORATE SOURCE: Pharmaceutical Products Division, Abbott Laboratories,

SOURCE: Abbott Park, IL, 60064, USA  
 Peptides: Chemistry, Structure and Biology,  
 Proceedings of the American Peptide Symposium, 14th,  
 Columbus, Ohio, June 18-23, 1995 (1996),  
 Meeting Date 1995, 663-664. Editor(s): Kaumaya,  
 Pravin T. P.; Hodges, Robert S. Mayflower Scientific:  
 Kingswinford, UK.  
 CODEN: 63NTAF

DOCUMENT TYPE: Conference

LANGUAGE: English

AB In an effort to improve the duration of action and the safety of LHRH antagonists we examd. conformational restrictions of the N-terminus of the know decapeptide antagonist, NACD2Nal-D4ClPhe-D3Pal-Ser-NMeTyr-DCit-Leu-Arg-Pro-DAla-NH2. Upon restriction of the N-terminus, either by N-methylation or by linking the acetyl group to the .alpha.-N-Me, we produced LHRH antagonists which in vitro were equally potent to the parent in inhibiting LH release, and in the castrate rat had longer duration of action. One of the analogs also had a better safety profile than the parent compd.

IT 184702-75-8

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
 (structure activity relations of LHRH antagonists conformationally restricted at the N-terminus)

L30 ANSWER 6 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1996:695912 HCAPLUS

DOCUMENT NUMBER: 126:14333

TITLE: Arginyl methylketones in the design of highly potent bivalent thrombin inhibitors

AUTHOR(S): Steinmetzer, T.; Rehse, P.; Zhu, B. Y.; Gibbs, B. F.; Lefebvre, J.; Cygler, M.; Konishi, Y.

CORPORATE SOURCE: Biotechnology Research Institute, National Research Council Canada, Montreal, QC, H4P 2R2, Can.

SOURCE: Peptides: Chemistry, Structure and Biology,  
 Proceedings of the American Peptide Symposium, 14th,  
 Columbus, Ohio, June 18-23, 1995 (1996),  
 Meeting Date 1995, 356-357. Editor(s): Kaumaya,  
 Pravin T. P.; Hodges, Robert S. Mayflower Scientific:  
 Kingswinford, UK.  
 CODEN: 63NTAF

DOCUMENT TYPE: Conference

LANGUAGE: English

AB Synthetic inhibitors, which mimic the binding mode of hirudin to thrombin, have been previously developed. They are composed of an active site inhibitor segment, a fibrinogen recognition exosite inhibitor segment, and a linker connecting these parts. Arginyl methylketones derivs. were incorporated in the P1-P1' region of the active site inhibitor segment and enhanced the binding affinity of the inhibitors. The synthesis and inhibitory potency of new bivalent thrombin inhibitors is presented.

IT 183969-25-7P

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
 (arginyl methylketones in design of highly potent bivalent thrombin inhibitors)

L30 ANSWER 7 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

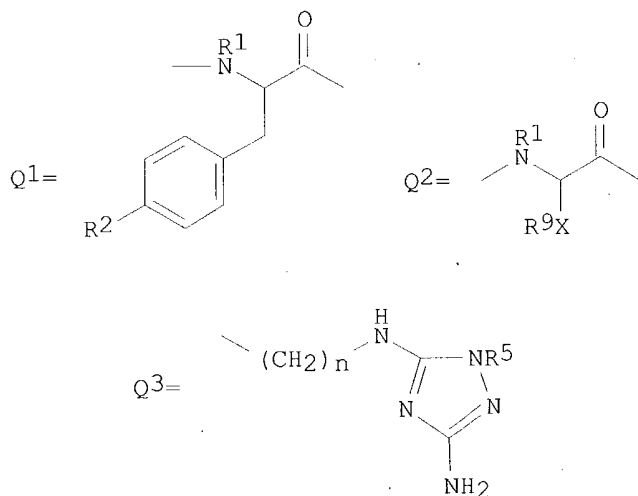
ACCESSION NUMBER: 1995:397087 HCAPLUS

DOCUMENT NUMBER: 122:161380

TITLE: Preparation of LHRH antagonists having modified amino acid residues at postions 5 and 6.

INVENTOR(S): Haviv, Fortuna; Greer, Jonathan; Swenson, Rolf E.;  
 Sauer, Daryl R.  
 PATENT ASSIGNEE(S): Abbott Laboratories, USA  
 SOURCE: PCT Int. Appl., 60 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9414841	A1	19940707	WO 1993-US12196	19931214 <--
W: CA, JP				
RW: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
CA 2136079	AA	19940707	CA 1993-2136079	19931214 <--
EP 683792	A1	19951129	EP 1994-905391	19931214 <--
EP 683792	B1	20010926		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, LU, NL, PT, SE				
JP 08505611	T2	19960618	JP 1993-515274	19931214 <--
AT 206136	E	20011015	AT 1994-905391	19931214
ES 2164096	T3	20020216	ES 1994-905391	19931214
US 5491217	A	19960213	US 1994-282411	19940728 <--
PRIORITY APPLN. INFO.:			US 1992-993202	A 19921218
			WO 1993-US12196	W 19931214
OTHER SOURCE(S):			MARPAT 122:161380	
GI				



AB A1B2C3D4E5F6G7H8I9J10 [A1 = N-acetyl-D-3-(2-naphthyl)alanyl, Ac-Sar, N-acetylazaglycyl, Ac-D-Phe, etc.; B1 = D-Phe, D-3-(4-chlorophenyl)alanyl, D-3-(2-naphthyl)alanyl, etc.; C3 = D-3-(3-pyridyl)alanyl, D-3-(2-thiazolyl)alanyl, etc.; D4 = Ser, N(R1)-substituted Ser; R1 = alkyl; E5 = Q1, Q2; R2 = NO2, CH2Cl, CH2OH, CH2N3, CH2CN, (CH2)mNR3R4, Q3, etc.; R3, R4 = H, alkyl, (substituted) Ph, PhCH2; NR3R4 = pyrrolidinyl, piperidinyl, morpholinyl, etc.; R5 = H, alkyl; m = 1,2; n = 0-2; X = 1,4-cyclohexylene, alkylene; R9 = (CH2)mNR3R4, Q3, etc.; F6 = D-Trp, D-3-(3-pyridyl)alanyl, D-Ser, Q1, etc.; G7 = Leu, N(R1)-substituted Leu, Val, cyclohexylalanyl, Ile, etc.; H8 = (.epsilon.-N-isopropyl)lysyl, N(R1)-substituted Arg; I9 = Pro, N(R1)-substituted Ala; J10 = NHet, D-Ala-NH2, Sar-NH2, D-Ser-NH2, etc.; with the proviso that when J = NHet, then I = Pro], were prepd. Thus, Ac-D-2Nal-D-Phe(4-Cl)-D-3Pal-Ser-

NMePhe(4-NO<sub>2</sub>)-D-Cit-Leu-Arg-Pro-D-Ala-NH<sub>2</sub>, [2Nal = 3-(2-naphthyl)alanyl, 3Pal = 3-(pyrid-2-yl)alanyl, Cit = citrullyl] prepd. using BOC-protected amino acids and methylbenzhydrylamine resin, antagonized LHRH with pA<sub>2</sub> = 11.26 using the methods of F. Haviv.

IT **161356-82-7P 161356-86-1P 161356-87-2P**  
**161356-91-8P 161356-92-9P 161356-93-0P**

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); SPN (Synthetic preparation); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)

(prepn. of LHRH antagonists having modified amino acid residues at positions 5 and 6)

L30 ANSWER 8 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1992:585141 HCAPLUS

DOCUMENT NUMBER: 117:185141

TITLE: Stabilization of the N-terminal residues of luteinizing hormone-releasing hormone agonists and the effect on pharmacokinetics

AUTHOR(S): Haviv, Fortuna; Fitzpatrick, Timothy D.; Nichols, Charles J.; Swenson, Rolf E.; Bush, Eugene N.; Diaz, Gilbert; Nguyen, A.; Nellans, Hugh N.; Hoffman, Daniel J.; et al.

CORPORATE SOURCE: Pharm. Prod. Div., Abbott Lab., Abbott Park, IL, 60064, USA

SOURCE: Journal of Medicinal Chemistry (1992), 35(21), 3890-4  
 CODEN: JMCMAR; ISSN: 0022-2623

DOCUMENT TYPE: Journal

LANGUAGE: English

AB To stabilize leuprolide, [D-Leu<sup>6</sup>,Pro<sup>9</sup>-NHet]LHRH (LHRH = LH-releasing hormone), against chymotrypsin and intestinal degrdn., several agonists of LHRH, e.g. [N-Me-Ser<sup>4</sup>,D-Leu<sup>6</sup>,Pro<sup>9</sup>-NHet]LHRH, [N-Ac-Sar<sup>1</sup>,D-Leu<sup>6</sup>,Pro<sup>9</sup>-NHet]LHRH (Sar = sarcosine), [Phe<sup>2</sup>,D-Trp,Pro<sup>9</sup>-NHet]LHRH, [N-MePhe<sup>2</sup>,D-Leu<sup>6</sup>,Pro<sup>9</sup>-NHet]LHRH, [Tyr(Me)<sup>3</sup>,D-Leu<sup>6</sup>,Pro<sup>9</sup>-NHet]LHRH, modified at positions 1, 2, or 3 and/or contg. N-.alpha.-Me at positions 1, 2, or 4, were synthesized by the solid-phase method. These agonist were tested in vitro for (a) rat pituitary LHRH receptor binding, (b) LH release from rat pituitary cells, (c) stability against chymotrypsin, and (d) stability against rat intestinal degrdn. The clearances of the compds. in the rat were detd. using a RIA. Complete stabilization against chymotrypsin (t<sub>1/2</sub>) and luminal degrdn. (T<sub>1/2</sub>) was achieved with substitution of NMe-Ser<sup>4</sup> in leuprolide; however, with an increase in clearance. Substitution with 1-Nal<sup>3</sup> (Nal = naphthylalanine) increased both t<sub>1/2</sub> and T<sub>1/2</sub>, while substitution with NAc-Sar<sup>1</sup> increased only T<sub>1/2</sub>. [NAcSar<sup>1</sup>,NMeSer<sup>4</sup>,D-Trp<sup>6</sup>,Pro<sup>9</sup>NH<sup>2</sup>]LHRH, the doubly stabilized analog, was tested in the rat by both i.v. and id administrations, and its bioavailabilities were measured. No significant improvement in id absorption over leuprolide was obsd.

IT **125323-88-8P 143399-02-4P**

RL: SPN (Synthetic preparation); PREP (Preparation)  
 (prepn., LH-releasing activity and stability of, to chymotrypsin and intestinal degrdn.)

L30 ANSWER 9 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1991:632885 HCAPLUS

DOCUMENT NUMBER: 115:232885

TITLE: Preparation of LHRH analogs

INVENTOR(S): Haviv, Fortuna; Greer, Jonathan

PATENT ASSIGNEE(S): Abbott Laboratories, USA

SOURCE: Eur. Pat. Appl., 79 pp.

CODEN: EPXXDW

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 413209	A1	19910220	EP 1990-114829	19900802 <--
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE				
JP 04505750	T2	19921008	JP 1989-502677	19890209 <--
US 5110904	A	19920505	US 1990-548512	19900710 <--
CA 2022444	AA	19910208	CA 1990-2022444	19900801 <--
CA 2022444	C	20020115		
NO 9003455	A	19910208	NO 1990-3455	19900806 <--
HU 54387	A2	19910228	HU 1990-4910	19900806 <--
AU 9060286	A1	19910207	AU 1990-60286	19900807 <--
JP 03101695	A2	19910426	JP 1990-209058	19900807 <--
US 5300492	A	19940405	US 1991-784204	19911028 <--
AU 9457892	A1	19940526	AU 1994-57892	19940317 <--
AU 672474	B2	19961003		

PRIORITY APPLN. INFO.:  
US 1989-390572 A 19890807  
US 1990-548512 A 19900710  
US 1988-154681 A2 19880210  
WO 1989-US528 W 19890209

OTHER SOURCE(S): MARPAT 115:232885

AB LHRH analogs A-B-C-D-E-F-G-H-I-J [A = amino acyl, e.g., L- or D-pyroglyutamyl, N-acetyl-L-prolyl, etc.; B = bond, amino acid residue, e.g., L- or D-Trp, etc.; C = amino acid residue, e.g., L- or D-Trp, D-Pro, etc.; D = amino acid residue, e.g., Pro, Pro(4-OH), etc.; E = amino acid residue, e.g., L-Tyr, L-Tyr(Me), etc.; F = amino acid residue; G = amino acid residue, e.g., L-Leu, L-Ile, etc., or F and G taken together are substituted gamma-lactam residue; H = NR1CH[(CH2)pR2]CO; R1 = H, Me, Et, Pr, Me2CH; R2 = (alkyl)amino(cyclohexyl), etc.; p = 1-4; I = imino acid or aliph. amino acid residue, e.g., L-Pro, L-MeAla, etc.; J = 1-pyrrolidinyl, 1-piperidinyl, 4-morpholinyl, or amino acid residue, e.g., D-alanylamide, etc.; with provisos] were prepd. Thus, H-(pyro)Glu-His-Trp-MeSer-Tyr-D-Leu-Leu-Arg-Pro-NHET (I) was prepd. using solid phase methods by sequential coupling of appropriate protected amino acids followed by deprotection and isolation as the trifluoroacetate salt. I.cntdot.CF3CO2H had an ED50 of 7.20 .mu.g/kg i.v. for LH release in castrated rats, compared to 100 .mu.g/kg for LHRH.

IT 125323-88-8P 125323-90-2P 135185-11-4P  
135185-12-5P 135185-14-7P 135185-31-8P  
135185-33-0P 135185-35-2P 135185-37-4P  
135185-39-6P 135185-59-0P 135185-61-4P

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. of, as LH-RH agonist)

L30 ANSWER 10 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1990:99259 HCAPLUS  
DOCUMENT NUMBER: 112:99259  
TITLE: Preparation of LH-RH analogs  
INVENTOR(S): Haviv, Fortuna; Greer, Jonathan  
PATENT ASSIGNEE(S): Abbott Laboratories, USA  
SOURCE: Eur. Pat. Appl., 39 pp.  
CODEN: EPXXDW

DOCUMENT TYPE: Patent  
LANGUAGE: English

FAMILY ACC. NUM. COUNT: 2

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 328090	A2	19890816	EP 1989-102208	19890209 <--
R: ES, GR				

WO 8907450 A1 19890824 WO 1989-US528 19890209 <--  
W: BE, CH, DE, FR, GB, IT, NL, SE  
EP 400065 A1 19901205 EP 1989-902885 19890209 <--  
EP 400065 B1 19970827  
R: BE, CH, DE, FR, GB, IT, LI, NL, SE  
JP 04505750 T2 19921008 JP 1989-502677 19890209 <--  
ES 2108684 T3 19980101 ES 1989-902885 19890209 <--  
US 5300492 A 19940405 US 1991-784204 19911028 <--

## PRIORITY APPLN. INFO.:

US 1988-154681 A 19880210  
WO 1989-US528 W 19890209  
US 1989-390572 B2 19890807  
US 1990-548512 A3 19900710

AB H-A-B-C-D-E-F-G-Q-I-J [A = L-pyroglutamyl, acetylprolyl, acetylphenylalanyl, etc.; B = His, Phe, naphthylalanyl, etc.; C = Trp, Phe, methylphenylalanyl, naphthylalanyl, etc.; D = Pro, benzylseryl, Ala, etc.; E = Tyr, O-methyltyrosyl, Phe, etc.; F = D-amino acid residue; G = Leu, Ile, Nle, etc.; Q = NR10CH[(CH2)pR9]CO; R9 = (substituted) aminocyclohexyl; R10 = H, alkyl; p = 1-4; I = (substituted) prolyl, methylalanyl, methylnorvalyl, etc.; J = 1-pyrrolidinyl, 1-piperidinyl, 4-morpholinyl, etc.], useful as LH-RH agonists and antagonists, were prepd. H-pyroGlu-His-Trp-N-Me-Ser-Tyr-D-Leu-Leu-Arg-Pro-NHEt, prepd. by the solid-phase method, had an ED50 of 7.20 .mu.g/kg i.v. in mice for half-maximal LH release.

IT **125323-85-5P 125323-88-8P 125323-90-2P**

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. of, as LHRH analog)

L30 ANSWER 11 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1986:424631 HCAPLUS  
DOCUMENT NUMBER: 105:24631  
TITLE: Therapeutic LHRH-related compounds  
PATENT ASSIGNEE(S): Biomeasure, Inc., USA  
SOURCE: Jpn. Kokai Tokkyo Koho, 6 pp.  
CODEN: JKXXAF  
DOCUMENT TYPE: Patent  
LANGUAGE: Japanese  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
JP 61007297	A2	19860113	JP 1984-252799	19841129 <--
US 4632979	A	19861230	US 1984-621673	19840618 <--
DK 8406150	A	19851219	DK 1984-6150	19841220 <--
EP 171477	A2	19860219	EP 1984-309021	19841221 <--
EP 171477	A3	19871014		
R: AT, BE, CH, DE, FR, GB, IT, LI, LU, NL, SE				
FI 8500193	A	19851219	FI 1985-193	19850116 <--
NO 8502439	A	19851219	NO 1985-2439	19850617 <--
AU 8543764	A1	19860102	AU 1985-43764	19850618 <--
ZA 8504590	A	19860226	ZA 1985-4590	19850618 <--
ES 544316	A1	19870101	ES 1985-544316	19850618 <--
CA 1257950	A1	19890725	CA 1985-484283	19850618 <--

## PRIORITY APPLN. INFO.:

US 1984-621673 19840618

GI

K-His-Trp-Ser-Tyr-M-Q-Arg-Pro-T I

AB The title compds. [I, K = N-acetyl-sarcosyl, pGlu; M = D-Phe, -Trp, -.beta.-naphthylalanyl, -4-X-Phe (X = OH, F, Cl, Br, Me); Q = Leu, Phe,

4-X-Phe, etc.; T = Gly-NH<sub>2</sub>, NHMe, NH<sub>2</sub>Et, NHPr], useful as anticarcinogenic agents (no data), were prepd. I (K = pGlu, M = D-Trp, Q = Leu, T = NH<sub>2</sub>Et) was among the prepd. compds. I may be administered i.m., s.c., p.o., etc.

IT **102865-90-7P 102865-91-8P 102907-98-2P**

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. of, as anticarcinogenic agent)

L30 ANSWER 12 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1984:523209 HCAPLUS

DOCUMENT NUMBER: 101:123209

TITLE: Design and synthesis of inhibitory analogs of LH-RH with increased antiovarulatory potency

AUTHOR(S): Mezo, I.; Seprodi, J.; Erchegeyi, J.; Horvath, A.;  
Nikolics, K.; Teplan, I.; Vigh, S.; Kovacs, M.;  
Flerko, B.; et al.

CORPORATE SOURCE: Med. Sch., Semmelweis Univ., Budapest, H-1444, Hung.

SOURCE: Acta Chimica Hungarica (1984), 116(2),  
173-87

CODEN: ACHUDC; ISSN: 0231-3146

DOCUMENT TYPE: Journal

LANGUAGE: English

AB In a structure-activity relation study, LH-RH analogs, some of which were prepd., were tested for ovulation-inhibiting potency in rats. Amino acid substitutions that increased the antiovarulatory potency of LH-RH analogs were: (1) acetyl-D-tryptophan substitution in pos. 1, (2) 4-chloro-D-phenylalanine substitution in pos. 2, and (3) D-phenylalanine substitution in pos. 6. However, [acetyl-D-Trp1, (4-chloro)-D-Phe2, D-Trp3, D-Lys6, D-Ala10]-LH-RH [87565-51-3] was more potent than [D-pyro-Glu1, (4-chloro)-D-Phe3, D-Trp3, 6]-LH-RH [75851-09-1] or [D-pyro-Glu1, (4-chloro)-D-Phe2, D-Trp3, D-Lys6]-LH-RH [75851-06-8]. Thus, the effectiveness of a specific substituent is basically dependent on the structure of the whole mol.

IT **86578-04-3**

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(ovulation-inhibiting activity of, structure in relation to)

L30 ANSWER 13 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1983:464462 HCAPLUS

DOCUMENT NUMBER: 99:64462

TITLE: The role of N-acyl groups in the inhibitory activity of LH-RH analogs

AUTHOR(S): Mezo, Imre; Seprodi, Janos; Erchegeyi, Judit; Teplan, Istvan; Kovacs, Magdolna; Flerko, Bela

CORPORATE SOURCE: Med. Sch., Semmelweis Univ., Budapest, Hung.

SOURCE: Peptides (New York, NY, United States) (1983), 4(2), 149-51

CODEN: PPTDD5; ISSN: 0196-9781

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Inhibitory analogs of LH-RH [9034-40-6] were prepd. with formyl-D-Trp1, acetyl-D-Trp1, valeryl-D-Trp1, tartaryl-D-Trp1, diacetyltartaryl-D-Trp1, acetyl-Gly1, and acetyl-Sar1 successively replacing the position 1 in the analog [D-Trp1, D-p-Cl-Phe2, D-Trp3, D-Phe6, D-Ala10]-LH-RH [86578-05-4]. The formyl-D-Trp1 and acetyl-D-Trp1 analogs yielded 100% blockade of ovulation at the 10 .mu.g dose; the others were less potent and inhibited ovulation at the 50 .mu.g dose. The inhibitory potency seems to correlate with the polarity of the acyl group.

IT **86578-04-3**

RL: BIOL (Biological study)  
(ovulation-inhibiting action of, structure in relation to)

L30 ANSWER 14 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1982:174580 HCAPLUS

DOCUMENT NUMBER: 96:174580

TITLE: Evolution of design and achievement of inhibitors of the luteinizing hormone-releasing hormone as inhibitors of ovulation

AUTHOR(S): Folkers, Karl; Humphries, John; Bowers, Cyril Y.

CORPORATE SOURCE: Inst. Biomed. Res., Univ. Texas, Austin, TX, 78712, USA

SOURCE: Zeitschrift fuer Naturforschung, Teil B: Anorganische Chemie, Organische Chemie (1982), 37B(2), 246-59

CODEN: ZNBAD2; ISSN: 0340-5087

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Structure-activity relations of LH-RH [9034-40-6] analogs as inhibitors of LH [9002-67-9] release and ovulation in rats and rhesus monkeys were studied. Inhibitory activities for >100 peptides are given. However, some analogs, e.g. [D-Phe2,Ala4,D-Phe6]-LH-RH [81419-23-0] (100 .mu.g) released LH and FSH [9002-68-0] at a ratio of LH/FSH greater than that induced by LH-RH. [D-Phe2,Pro3,D-Phe6]-LH-RH [64789-67-9] (6 S.c. injections of 50 mg every 8 h) inhibited ovulation and the action of endogenous LH-RH in cycling rhesus monkeys. Infusion of [D-Phe2,Pro3,D-Trp6]LH-RH [60961-52-6] (375 .mu.g/day for 4 days) from a s.c. implanted minipump inhibited ovulation in cycling female rats and inhibited LH release in castrated male rats. Infusion of LH-RH (375 .mu.g/day, 4 days) and [D-Ala6,de-Gly10]-LH-RH EtNH2 [52435-06-0] (6 .mu.g/day, 4 days) blocked uterine implantation sites of mated rats. Antagonist analogs with 3-proline and 3-leucine residues did not block the implantation sites indicating a difference in mechanism of contraception for agonists and antagonists of LH-RH. Solid phase synthesis of the peptides is also discussed.

IT 81419-14-9

RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(ovulation-inhibiting activity of, structure in relation to)

L30 ANSWER 15 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1978:437043 HCAPLUS

DOCUMENT NUMBER: 89:37043

TITLE: Pituitary receptor binding activity of active, inactive, superactive and inhibitory analogs of gonadotropin-releasing hormone

AUTHOR(S): Heber, D.; Odell, W. D.

CORPORATE SOURCE: Dep. Med., Univ. California Sch. Med., Torrance, CA, USA

SOURCE: Biochemical and Biophysical Research Communications (1978), 82(1), 67-73

CODEN: BBRCA9; ISSN: 0006-291X

DOCUMENT TYPE: Journal

LANGUAGE: English

AB In order to further characterize the gonadotropin-releasing hormone [33515-09-2] receptor, 20 synthetic analogs with known bioactivity were tested in a radioreceptor assay. In vivo biol. activity correlated with high affinity receptor binding but not with low affinity binding. Inhibitory analogs with no in vivo biol. activity and weak antagonistic properties did not bind, whereas in vivo active or superactive analogs bound to high affinity receptors. Apparently, the high affinity gonadotropin-releasing hormone receptor binds only biol. active gonadotropin-releasing hormone-like peptides and this binding may be the initial step in gonadotropin-releasing hormone actions at the pituitary level.



IT 62577-30-4

RL: PROC (Process)

(receptor binding of, in pituitary gland)

L30 ANSWER 16 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1977:155951 HCAPLUS

DOCUMENT NUMBER: 86:155951

TITLE: Synthesis and biological activity of position 1  
analogs of LH-RHAUTHOR(S): Nikolics, Karoly; Coy, David H.; Vilchez-Martinez,  
Jesus A.; Coy, Esther J.; Schally, Andrew V.

CORPORATE SOURCE: Sch. Med., Tulane Univ., New Orleans, LA, USA

SOURCE: International Journal of Peptide & Protein Research (  
1977), 9(1), 57-62

CODEN: IJPPC3; ISSN: 0367-8377

DOCUMENT TYPE: Journal

LANGUAGE: English

AB R-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH<sub>2</sub> [R = HCONMeCH<sub>2</sub>CO (I), AcNMeCH<sub>2</sub>CO (II), 2-pyrrolidin-4-ylcarbonyl (III), N-methyl-2-pyrrolidin-4-ylcarbonyl (IV), H-Hyp (V), cyclopentylcarbonyl (VI)], LH-releasing hormone (LH-RH) analogs, were prepd. by an automated solid-phase procedure using a benzhydrylamine resin. The LH- and FSH-releasing activities of I, II, III, and IV in immature male rats after s.c. administration were 64%, 72%, 19%, and 58%, resp., of that of LH-RH. V exhibited very low gonadotropin-releasing potency (0.001% of LH-RH), however VI was .apprx.1.4% as active as LH-RH despite the lack of functional groups present in pyroglutamic acid. Apparently, the .gamma.-carbonyl group of pyroglutamic acid is responsible for maintaining high levels of biol. activity in LH-RH.

IT 62577-30-4P

RL: SPN (Synthetic preparation); PREP (Preparation)

(prepn. and LH- and FSH-releasing activities of)

L30 ANSWER 17 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1977:5858 HCAPLUS

DOCUMENT NUMBER: 86:5858

TITLE: Peptide amides with LH-releasing hormone activity

INVENTOR(S): Baba, Yoshihiko; Okada, Yutaka; Horikoshi, Hiroyoshi;  
Yabe, Yuichiro

PATENT ASSIGNEE(S): Sankyo Co., Ltd., Japan

SOURCE: Jpn. Kokai Tokkyo Koho, 6 pp.

CODEN: JKXXAF

DOCUMENT TYPE: Patent

LANGUAGE: Japanese

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
JP 51088966	A2	19760804	JP 1975-13177	19750131 <--
JP 58010380	B4	19830225		

PRIORITY APPLN. INFO.: JP 1975-13177 19750131

AB Peptide amides RCONR<sub>1</sub>CH<sub>2</sub>CO-His-Trp-Ser-Tyr-Z-Leu-Arg-Pro-NHR<sub>2</sub> (I; R = C<sub>1</sub>-4 straight-chain or branched alkyl; R<sub>1</sub> = Me, Et; Z = D-.alpha.-alkylglycine residue where the alkyl = Me, Et, Pr, Bu; R<sub>2</sub> = Et, 2-mono- to trifluoroethyl, but not Z = D-Ala and R<sub>2</sub> = Et) or their salts were prepd. by acylating His-Trp-Ser-Tyr-Z-Leu-Arg-Pro-NHR<sub>2</sub> (II) with RCONR<sub>1</sub>CH<sub>2</sub>CO<sub>2</sub>H or their reactive derivs. Thus, II (Z = D-Ala, R<sub>2</sub> = CH<sub>2</sub>CF<sub>3</sub>), prepd. from 100 mg N-benzyloxycarbonyl deriv., was acylated with N-acetylsarcosine 2,3,5-trichlorophenyl ester in DMF at room temp. for 6 hr and fractionated with Sephadex G-25 to give 43 mg I (R = R<sub>1</sub> = Me, Z = D-Ala, R<sub>2</sub> = CH<sub>2</sub>CF<sub>3</sub>), which had higher activity than natural LH-releasing hormone in rats. Also prepd. were I (R = R<sub>1</sub> = Me, Z = D-Nle) where R<sub>2</sub> = Et, CH<sub>2</sub>CH<sub>2</sub>F, CH<sub>2</sub>CF<sub>3</sub>.

IT 58083-03-7P

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. and leuteinizing hormone releasing activity of)

IT 58083-04-8P 58083-05-9P 58141-55-2P

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. of)

L30 ANSWER 18 OF 18 HCAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 1976:60007 HCAPLUS

DOCUMENT NUMBER: 84:60007

TITLE: Luteinizing hormone-releasing factor analogs

INVENTOR(S): Baba, Yoshihiko; Okada, Yutaka; Horikoshi, Hiroyoshi;  
Yabe, Yuichiro

PATENT ASSIGNEE(S): Sankyo Co., Ltd., Japan

SOURCE: Ger. Offen., 21 pp.

CODEN: GWXXBX

DOCUMENT TYPE: Patent

LANGUAGE: German

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
DE 2517512	A1	19751113	DE 1975-2517512	19750421 <--
JP 50142563	A2	19751117	JP 1974-48174	19740426 <--
US 4010149	A	19770301	US 1975-565362	19750407 <--
GB 1473795	A	19770518	GB 1975-16046	19750418 <--
FR 2322610	A1	19770401	FR 1975-13042	19750425 <--
FR 2322610	B1	19790608		
CH 595328	A	19780215	CH 1975-5313	19750425 <--

PRIORITY APPLN. INFO.:

JP 1974-48174 19740426

AB AcNMeCH<sub>2</sub>CO-His-Trp-Ser-Tyr-X-Leu-Arg-Pro-NHCH<sub>2</sub>R (X = D-Ala, R = Me (I), CF<sub>3</sub>; X = D-Nle, R = Me, CH<sub>2</sub>F (II), CF<sub>3</sub>) were prepd. by condensation of the activated ester of AcNMeCH<sub>2</sub>CO<sub>2</sub>H with the appropriately blocked octapeptide amide. The ED<sub>50</sub> for inducing ovulation in rats was 0.018 and 0.0125 .mu.g for I and II, resp., in comparison to 0.134 .mu.g for LH releasing hormone.

IT 58083-02-6P 58083-04-8P

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. and ovulation inducing activity of)

IT 58083-03-7P 58083-05-9P 58141-55-2P

RL: SPN (Synthetic preparation); PREP (Preparation)  
(prepn. of)

=&gt; select hit rn l30 1-18

E18 THROUGH E57 ASSIGNED

=&gt; fil reg

FILE 'REGISTRY' ENTERED AT 16:26:01 ON 28 SEP 2003

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file  
provided by InfoChem.

STRUCTURE FILE UPDATES: 26 SEP 2003 HIGHEST RN 593958-55-5

DICTIONARY FILE UPDATES: 26 SEP 2003 HIGHEST RN 593958-55-5

TSCA INFORMATION NOW CURRENT THROUGH JULY 14, 2003

Please note that search-term pricing does apply when

conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:  
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=>  
=>

=> d his 131 1-40  
 '1-40' IS NOT VALID HERE

=> d his 131

(FILE 'HCAPLUS' ENTERED AT 16:25:28 ON 28 SEP 2003)  
 SELECT HIT RN L30 1-18

FILE 'REGISTRY' ENTERED AT 16:26:01 ON 28 SEP 2003  
 L31 40 S E18-E57

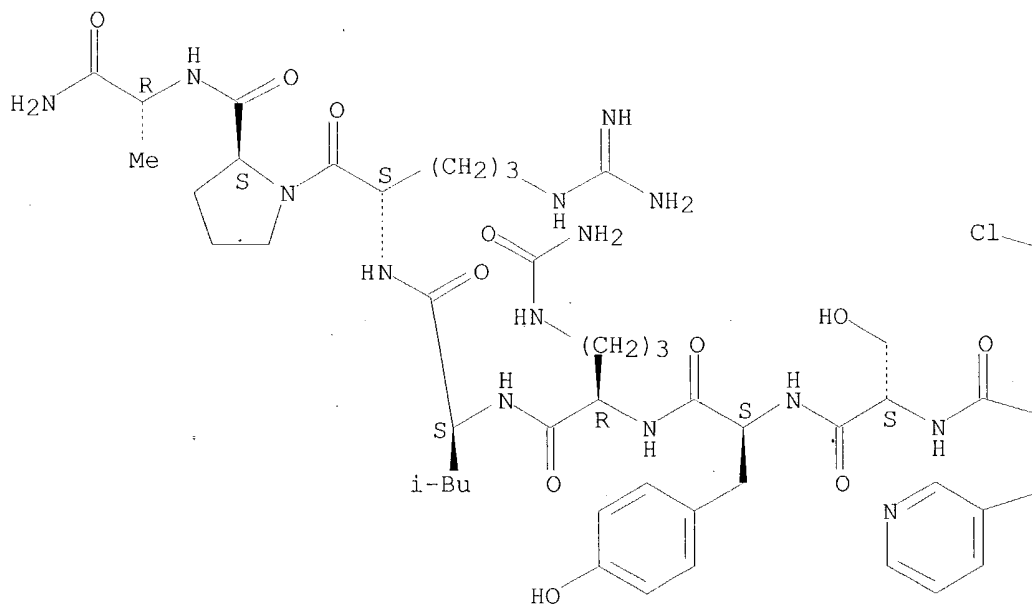
=> d ide can 131 1-40

L31 ANSWER 1 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **258332-86-4** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-(1-naphthalenylmethyl)glycyl-4-chloro-D-phenylalanyl-3-(3-pyridinyl)-D-alanyl-L-seryl-L-tyrosyl-N5-(aminocarbonyl)-D-ornithyl-L-leucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C70 H92 Cl N17 O14  
 SR CA  
 LC STN Files: CA, CAPLUS

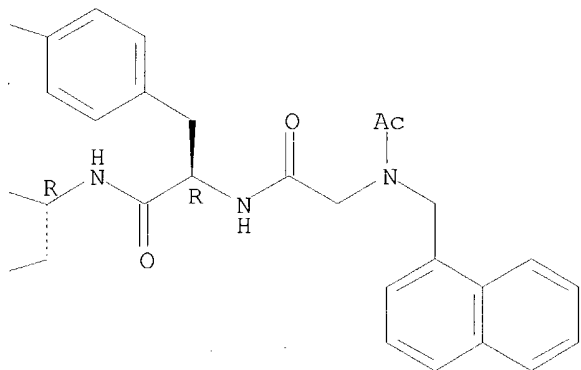
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

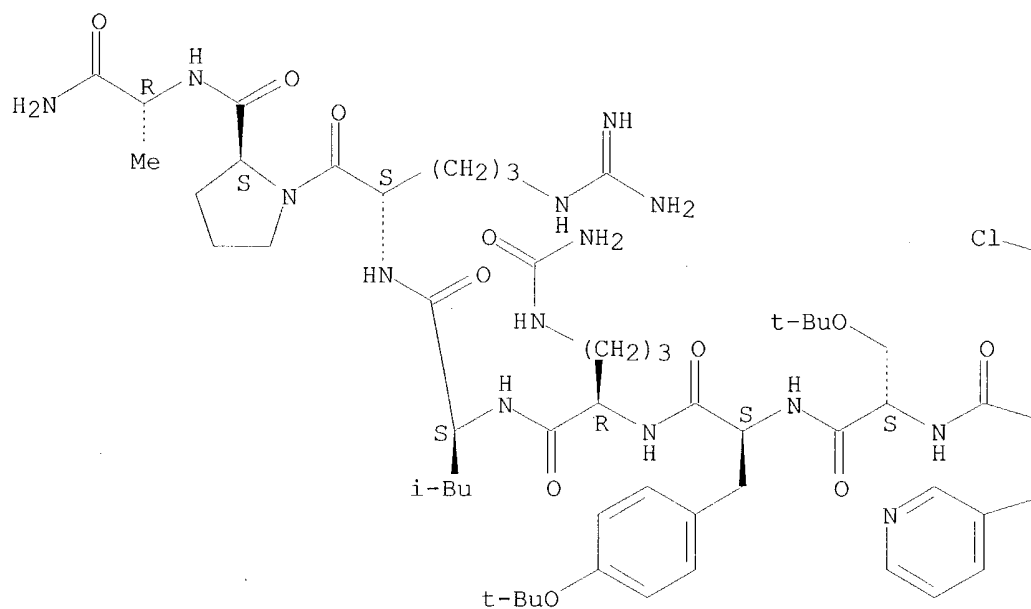
REFERENCE . 1: 132:152142

L31 ANSWER 2 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **258332-71-7** REGISTRY  
CN D-Alaninamide, N-acetyl-N-(1-naphthalenylmethyl)glycyl-4-chloro-D-phenylalanyl-3-(3-pyridinyl)-D-alanyl-O-(1,1-dimethylethyl)-L-seryl-O-(1,1-dimethylethyl)-L-tyrosyl-N5-(aminocarbonyl)-D-ornithyl-L-leucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C78 H108 Cl N17 O14  
SR CA  
LC STN Files: CA, CAPLUS

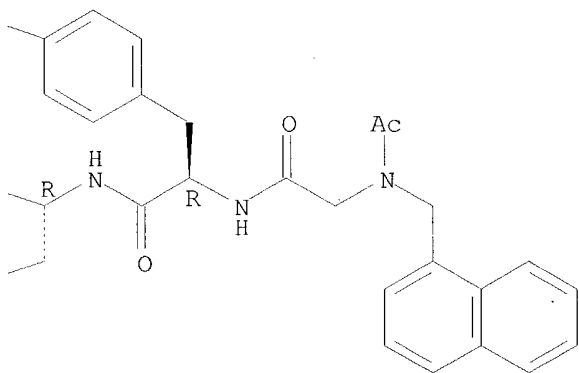
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:152142

L31 ANSWER 3 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 186837-31-0 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(2-naphthalenyl)-D-alanyl-L-seryl-L-tyrosyl-3-[1-(1-methylethyl)pyridinium-3-

yl]-L-alanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, salt with  
trifluoroacetic acid (1:1), mono(trifluoroacetate) (salt) (9CI) (CA INDEX  
NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C73 H99 Cl N13 O13 . C2 H F3 O2 . C2 F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER, USPAT2, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

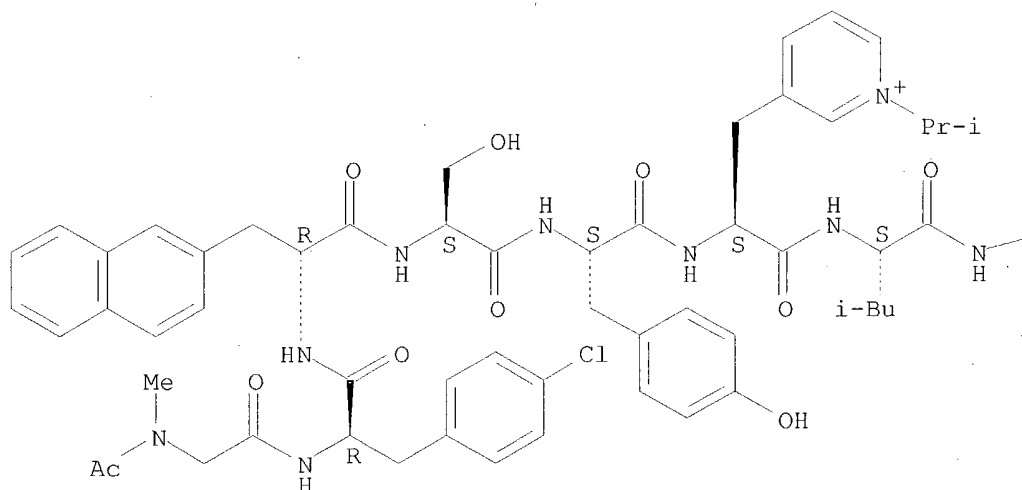
CRN 186837-30-9

CMF C73 H99 Cl N13 O13

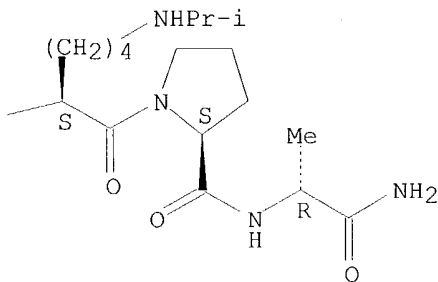
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

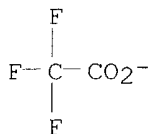


PAGE 1-B



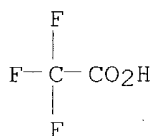
CM 2

CRN 14477-72-6  
CMF C2 F3 O2



CM 3

CRN 76-05-1  
CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:152828

L31 ANSWER 4 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **186837-28-5** REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(2-naphthalenyl)-D-alanyl-L-seryl-L-tyrosyl-3-[1-(1-methylethyl)pyridinium-3-yl]-D-alanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, salt with trifluoroacetic acid (1:1), mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C73 H99 Cl N13 O13 . C2 H F3 O2 . C2 F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER, USPAT2, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

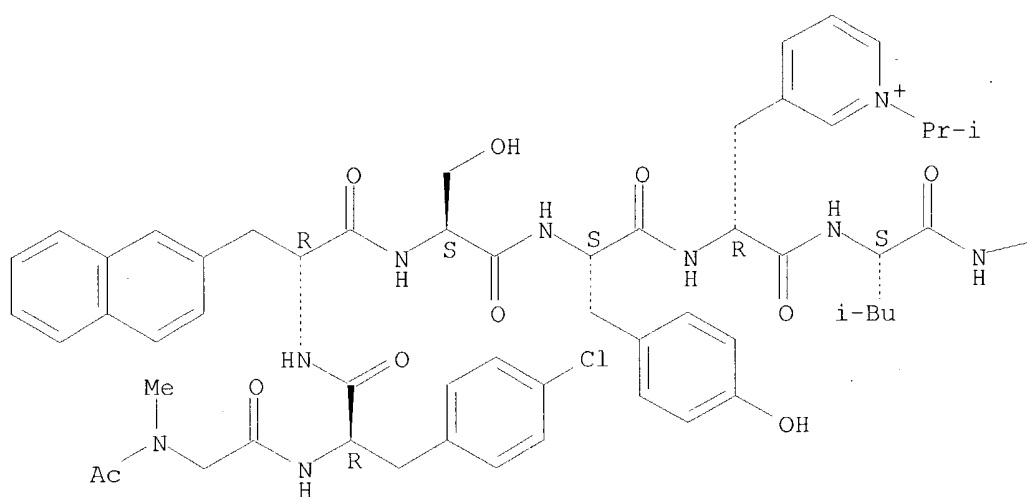
CM 1

CRN 186837-27-4  
CMF C73 H99 Cl N13 O13

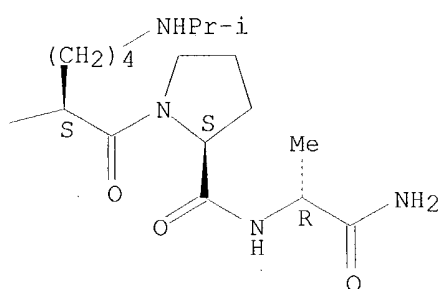
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



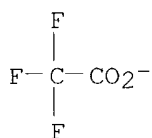
PAGE 1-B



CM 2

CRN 14477-72-6

CMF C2 F3 O2

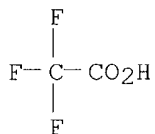


CM 3

CRN 76-05-1

CMF C2 H F3 O2





1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:152828

L31 ANSWER 5 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **186837-18-3** REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-L-alanyl-L-seryl-L-tyrosyl-3-[1-(phenylmethyl)pyridinium-3-yl]-D-alanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, salt with trifluoroacetic acid (1:1), mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C77 H99 Cl N13 O13 . C2 H F3 O2 . C2 F3 O2  
SR CA  
LC STN Files: CA, CAPLUS, TOXCENTER, USPAT2, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

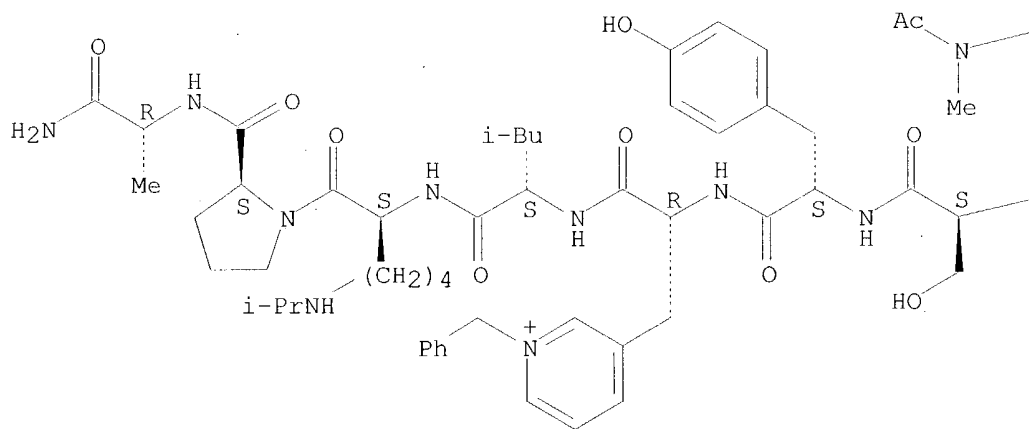
CM 1

CRN 186837-17-2  
CMF C77 H99 Cl N13 O13

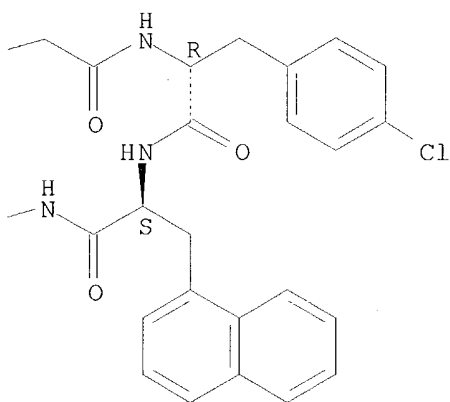
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



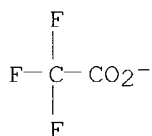
PAGE 1-B



CM 2

CRN 14477-72-6

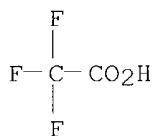
CMF C2 F3 O2



CM 3

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:152828

L31 ANSWER 6 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 186837-16-1 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(2-naphthalenyl)-D-alanyl-L-seryl-L-tyrosyl-3-[1-(phenylmethyl)pyridinium-3-yl]-D-alanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, salt with trifluoroacetic acid (1:1), mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C77 H99 Cl N13 O13 . C2 H F3 O2 . C2 F3 O2

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPAT2, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

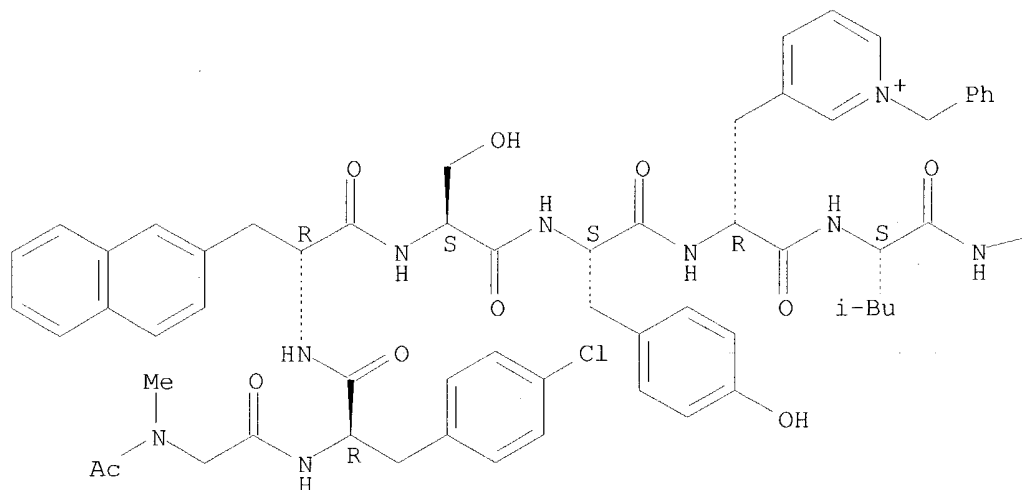
CRN 186837-15-0

CMF C77 H99 C1 N13 O13

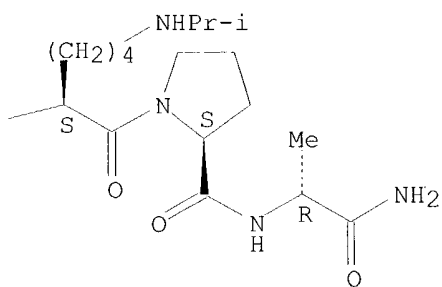
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



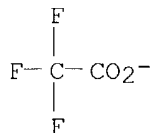
PAGE 1-B



CM 2

CRN 14477-72-6

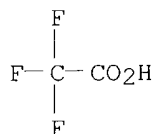
CMF C2 F3 O2



CM 3

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:152828

L31 ANSWER 7 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 186836-17-9 REGISTRY

CN D-Alaninamide; N-acetyl-N-methyl-3-(2-naphthalenyl)-D-alanyl-4-chloro-D-phenylalanyl-3-(3-pyridinyl)-D-alanyl-L-seryl-L-tyrosyl-N5-(aminocarbonyl)-D-ornithyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, mono(trifluoroacetate) (salt) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C74 H100 Cl N15 O14 . C2 H F3 O2

SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPAT2, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

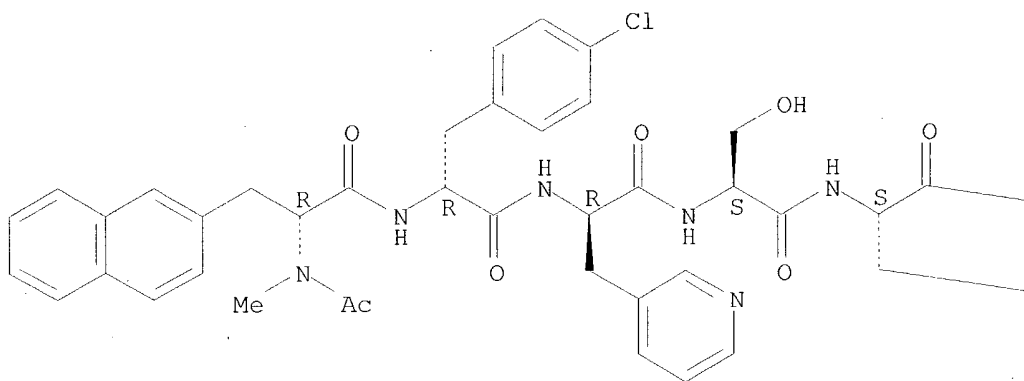
CRN 186836-16-8

CMF C74 H100 Cl N15 O14

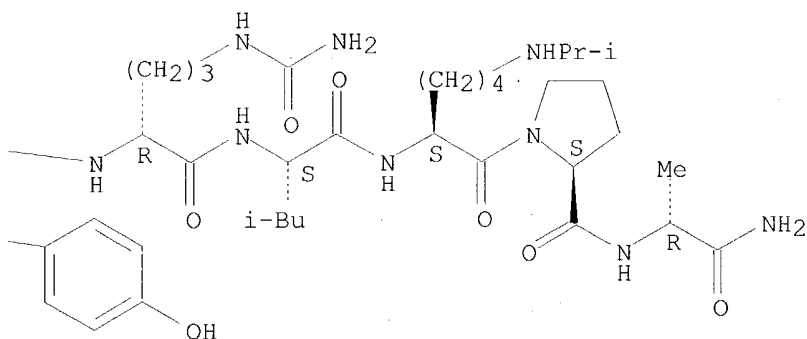
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



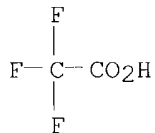
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:152828

L31 ANSWER 8 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 184702-75-8 REGISTRY

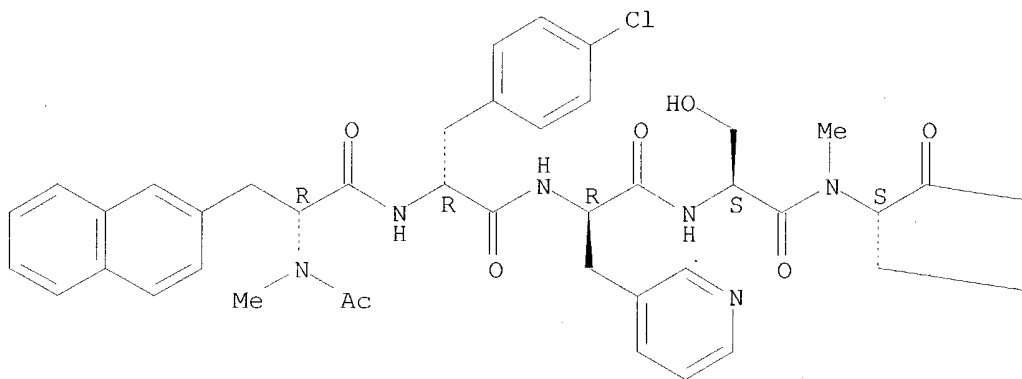
CN D-Alaninamide, N-acetyl-N-methyl-3-(2-naphthalenyl)-D-alanyl-4-chloro-D-phenylalanyl-3-(3-pyridinyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-N5-(aminocarbonyl)-D-ornithyl-L-leucyl-L-arginyl-L-prolyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C72 H96 Cl N17 O14  
 SR CA  
 LC STN Files: CA, CAPLUS

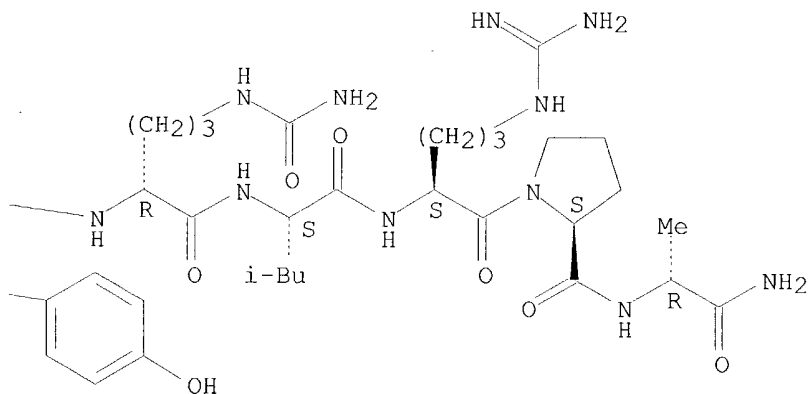
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 126:42251

L31 ANSWER 9 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **183969-25-7** REGISTRY

CN L-Aspartic acid, N-acetyl-N-[(3S)-6-[(aminoiminomethyl)amino]-3-[(3-cyclohexyl-D-alanyl-L-prolyl)amino]-2-oxohexyl]glycylglycylglycylglycylglycyl-L-.alpha.-aspartyl-L-tyrosyl-L-.alpha.-glutamyl-L-prolyl-L-isoleucyl-L-prolyl-L-.alpha.-glutamyl-L-.alpha.-glutamyl-L-tyrosyl-3-cyclohexyl-L-alanyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN P 536

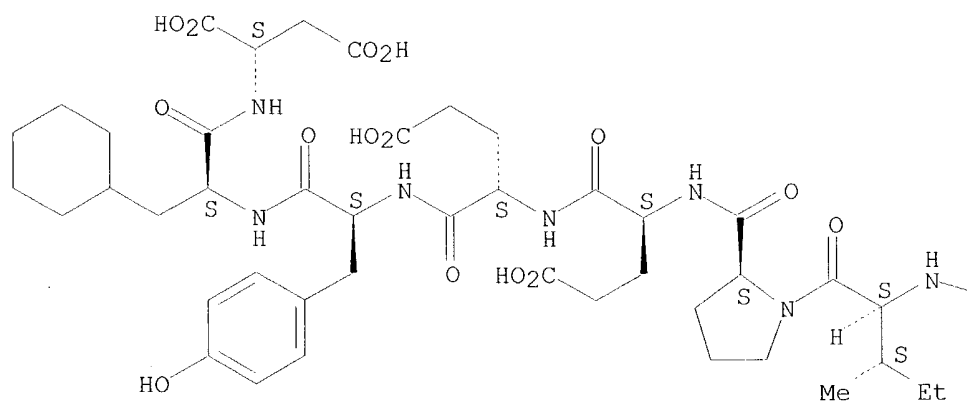
FS PROTEIN SEQUENCE; STEREOSEARCH

MF C99 H144 N22 O33

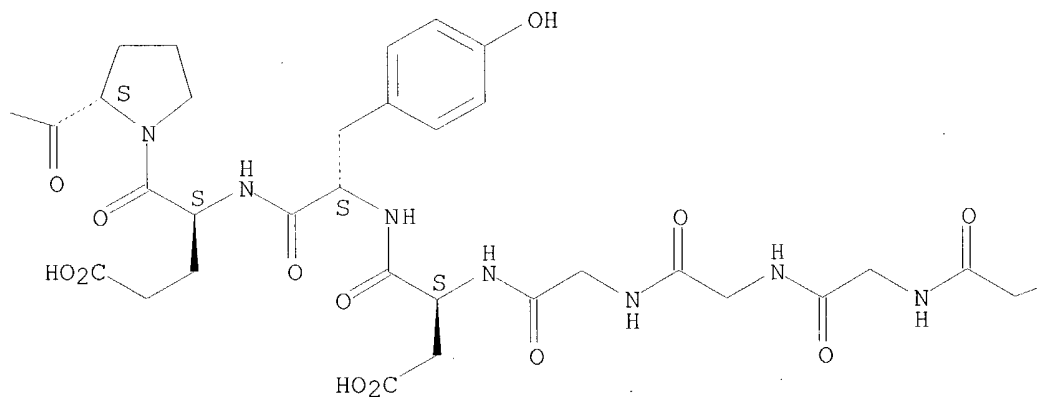
SR CA  
LC STN Files: CA, CAPLUS, USPATFULL

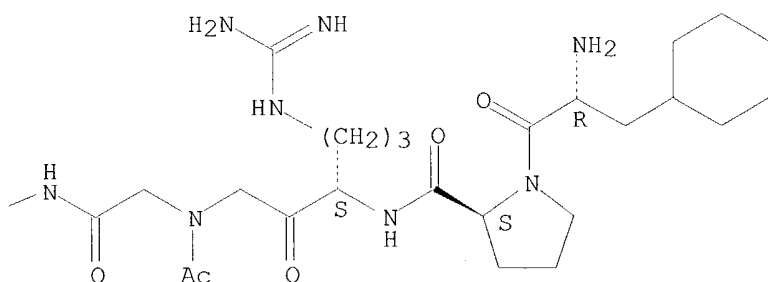
Absolute stereochemistry.

PAGE 1-A



PAGE 1-B





3 REFERENCES IN FILE CA (1907 TO DATE)

3 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 132:322148

REFERENCE 2: 131:237478

REFERENCE 3: 126:14333

L31 ANSWER 10 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **161356-93-0** REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-benzo[b]thien-4-yl-D-alanyl-L-seryl-N-methyl-L-tyrosyl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-D-phenylalanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C72 H96 Cl N17 O13 S

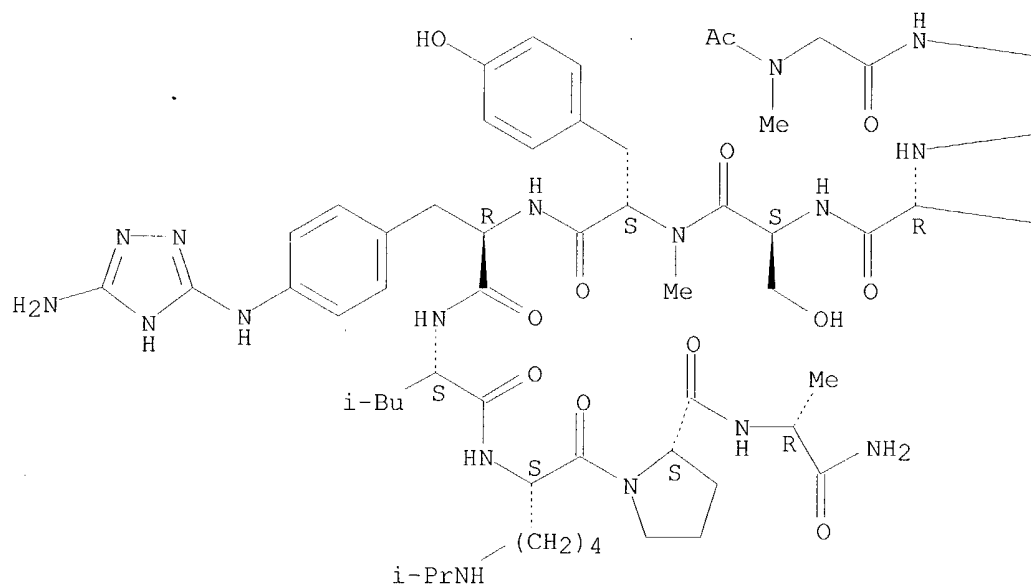
SR CA

LC STN Files: CA, CAPLUS, USPATFULL

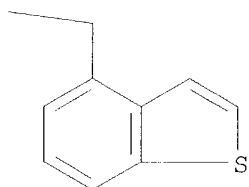
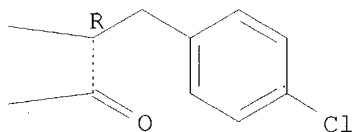
Absolute stereochemistry.



PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

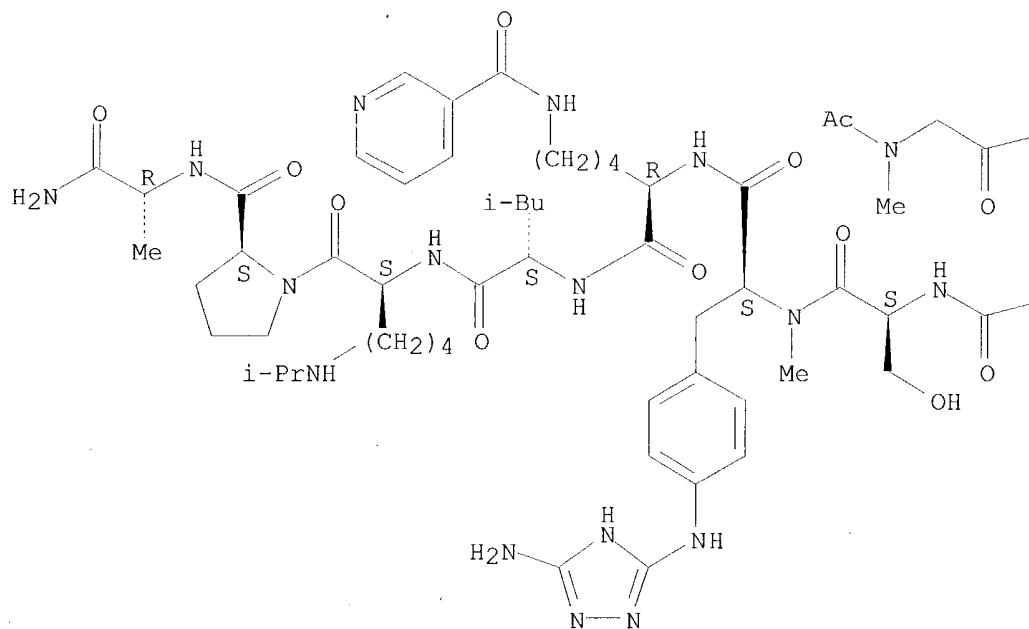
REFERENCE 1: 122:161380

L31 ANSWER 11 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **161356-92-9** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-benzo[b]thien-4-yl-D-alanyl-L-seryl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-N-methyl-L-phenylalanyl-N6-(3-pyridinylcarbonyl)-D-lysyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C75 H102 Cl N19 O13 S  
 SR CA  
 LC STN Files: CA, CAPLUS, USPATFULL

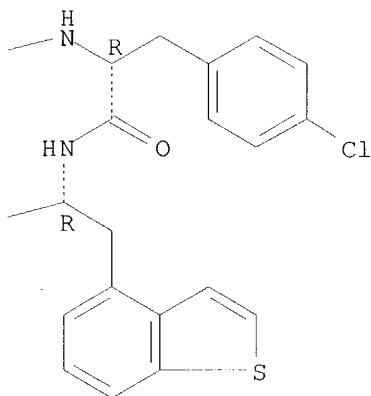
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 122:161380

L31 ANSWER 12 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 161356-91-8 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-benzo[b]thien-4-yl-D-alanyl-L-seryl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-N-methyl-L-phenylalanyl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-

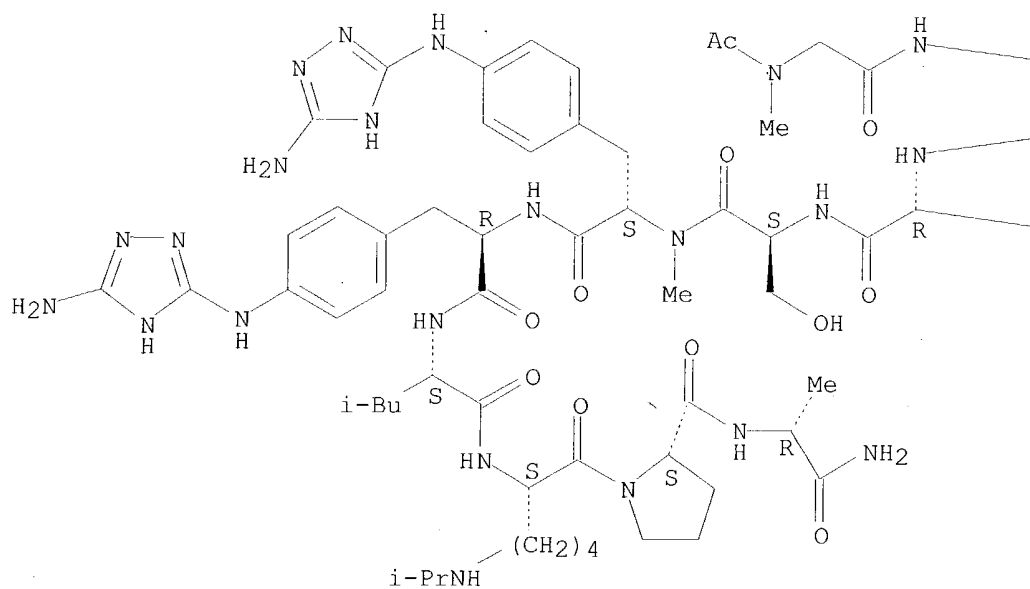
D-phenylalanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA  
INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C74 H99 Cl N22 O12 S  
SR CA  
LC STN Files: CA, CAPLUS, USPATFULL

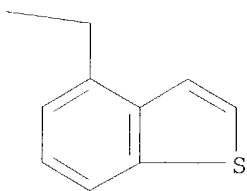
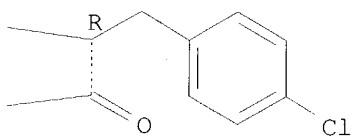
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 122:161380

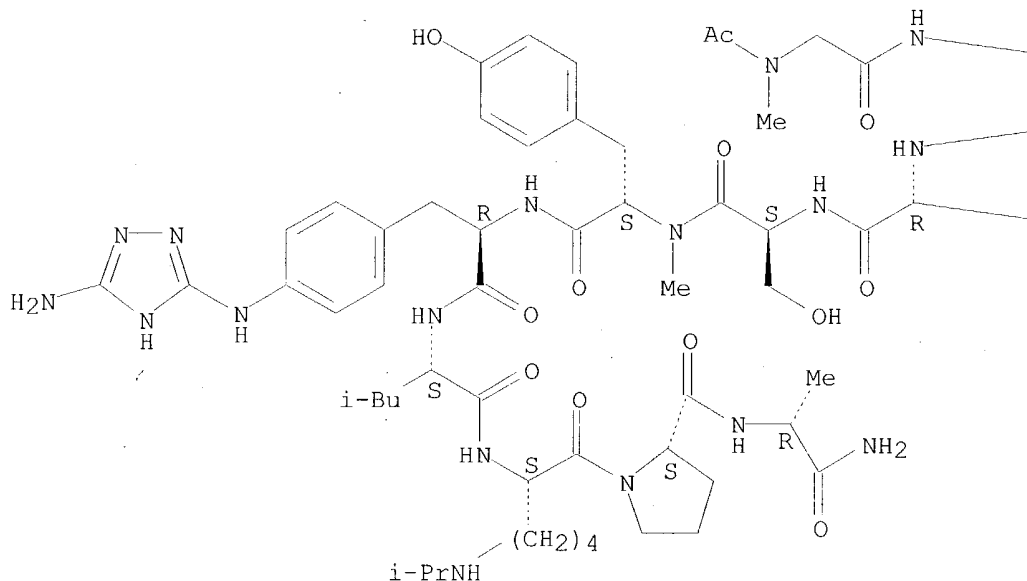
L31 ANSWER 13 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN 161356-87-2 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-D-phenylalanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C74 H98 Cl N17 O13  
 SR CA  
 LC STN Files: CA, CAPLUS, USPATFULL

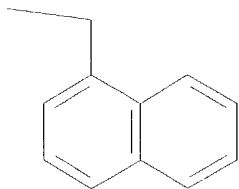
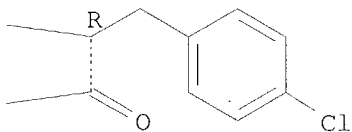
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

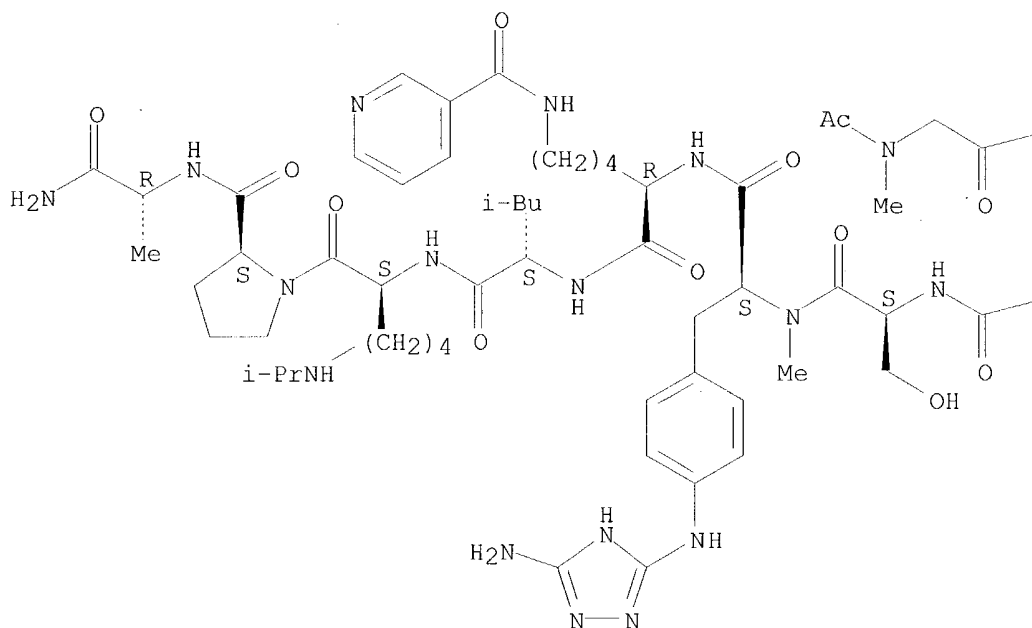
REFERENCE 1: 122:161380

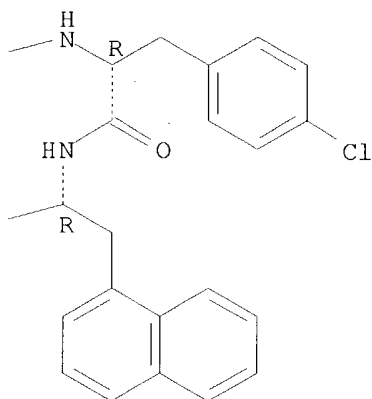
L31 ANSWER 14 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **161356-86-1** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-N-methyl-L-phenylalanyl-N6-(3-pyridinylcarbonyl)-D-lysyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C77 H104 Cl N19 O13  
 SR CA  
 LC STN Files: CA, CAPLUS, USPATFULL

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A





1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

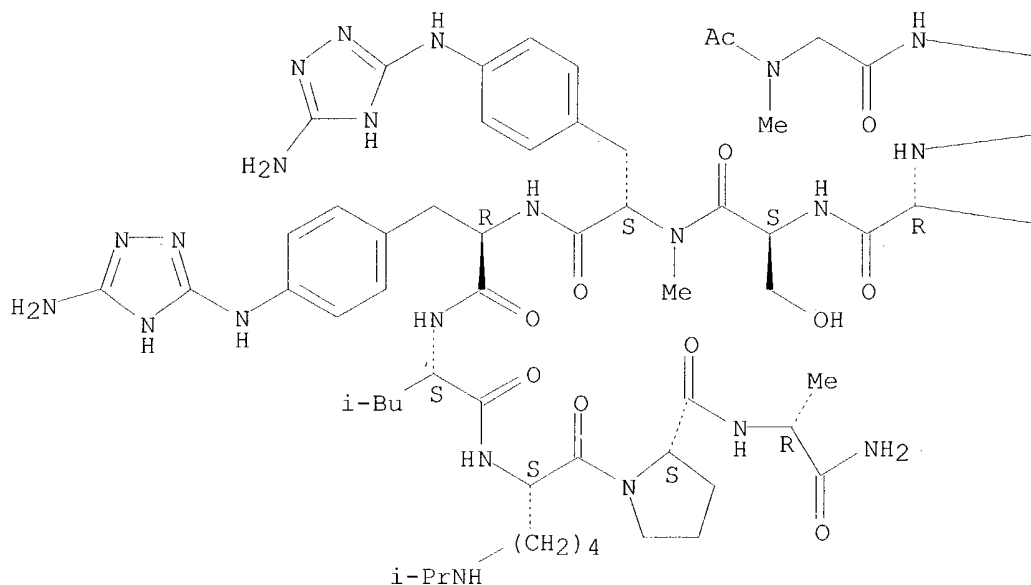
REFERENCE 1: 122:161380

L31 ANSWER 15 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **161356-82-7** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-N-methyl-L-phenylalanyl-4-[(5-amino-1H-1,2,4-triazol-3-yl)amino]-D-phenylalanyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C76 H101 Cl N22 O12  
 SR CA  
 LC STN Files: CA, CAPLUS, USPATFULL

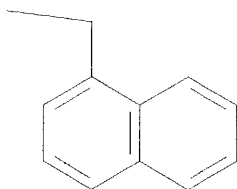
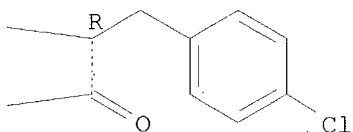
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 122:161380

L31 ANSWER 16 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 143399-02-4 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-methylglycine)-6-D-leucine-9-(N-ethyl-L-prolinamide)-10-deglycinamide- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-methylglycine)-6-D-leucine-9-(N-ethyl-L-prolinamide)-10-deglycinamide-

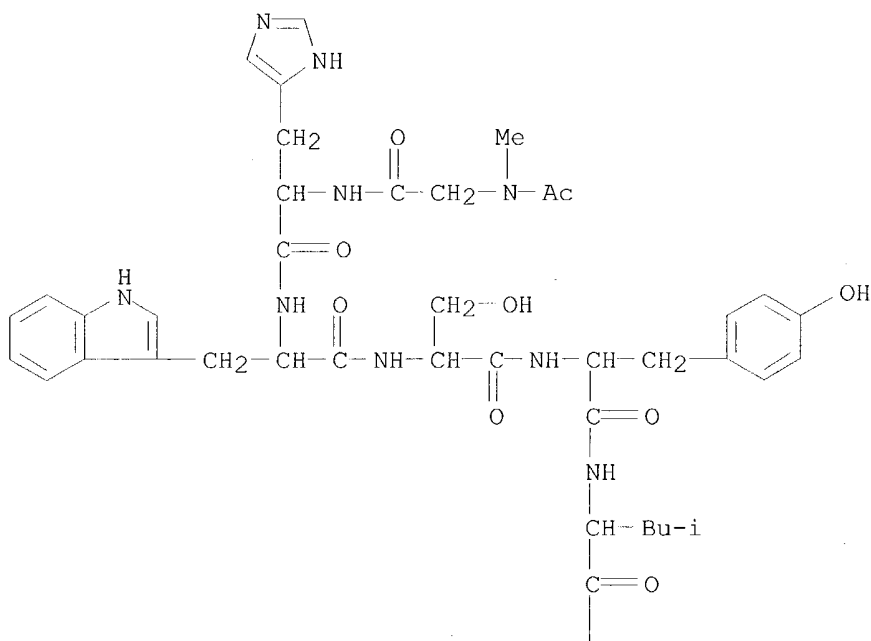
FS PROTEIN SEQUENCE

MF C59 H86 N16 O12

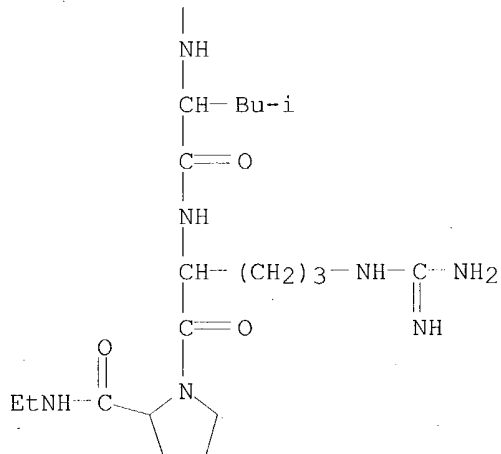
SR CA

LC STN Files: CA, CAPLUS

PAGE 1-A



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 117:185141

L31 ANSWER 17 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **135185-61-4** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-benzo[b]thien-3-yl-D-alanyl-L-seryl-N-methyl-L-tyrosyl-N6-(3-pyridinylcarbonyl)-D-lysyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH



MF C73 H99 Cl N14 O14 S . x C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

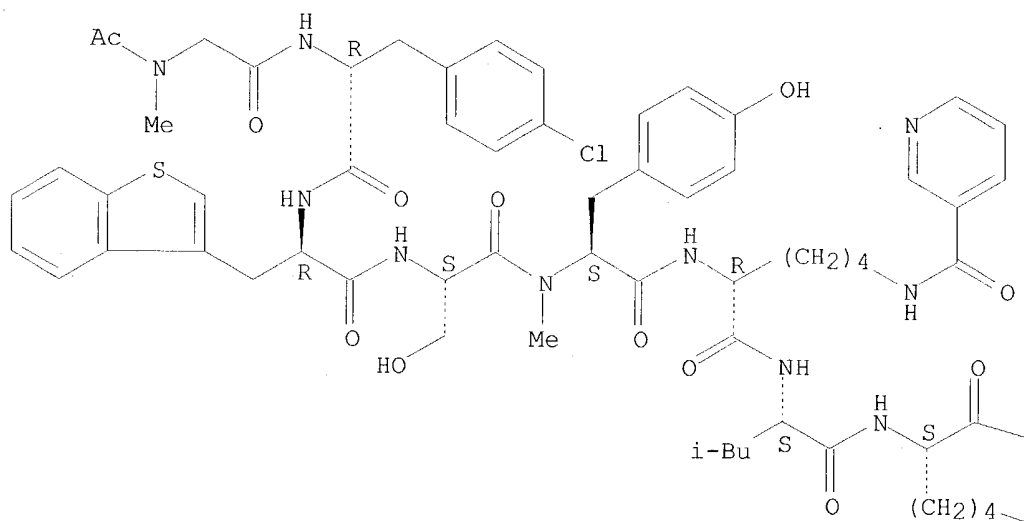
CRN 135185-60-3

CMF C73 H99 Cl N14 O14 S

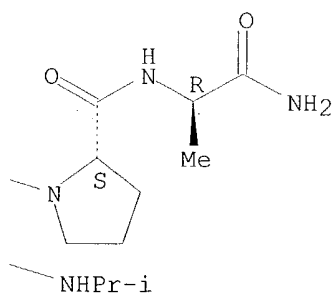
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

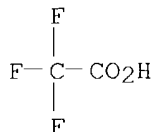
PAGE 1-A



PAGE 1-B



CM 2

CRN 76-05-1  
CMF C2 H F3 O21 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 18 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **135185-59-0** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-N6-(3-pyridinylcarbonyl)-D-lysyl-L-leucyl-N6-(1-methylethyl)-L-lysyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C75 H101 Cl N14 O14 . x C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

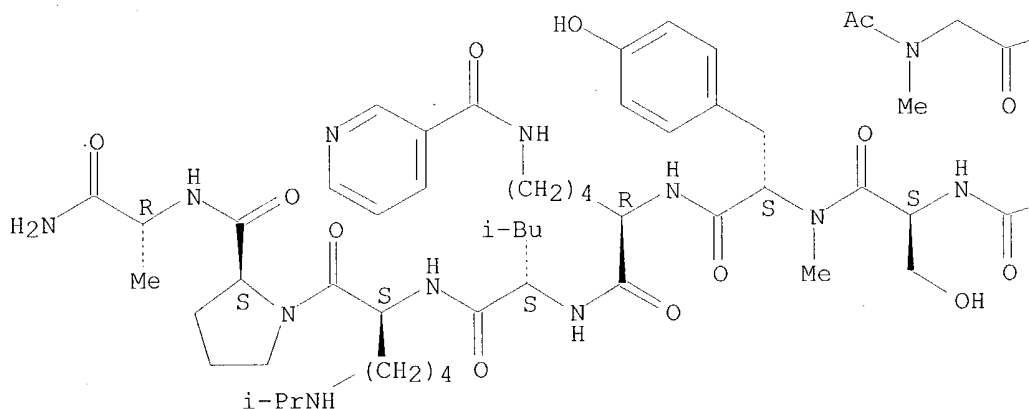
CM 1

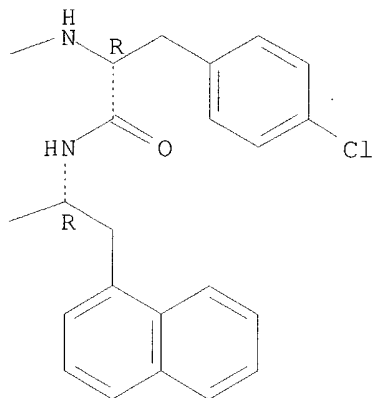
CRN 135185-58-9  
CMF C75 H101 Cl N14 O14

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

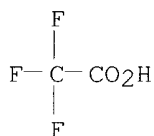
PAGE 1-A





CM 2

CRN 76-05-1  
 CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 19 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
 RN **135185-39-6** REGISTRY  
 CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-L-alanyl-L-seryl-N-methyl-L-tyrosyl-D-lysyl-L-leucyl-L-arginyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C66 H92 Cl N15 O13 . x C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

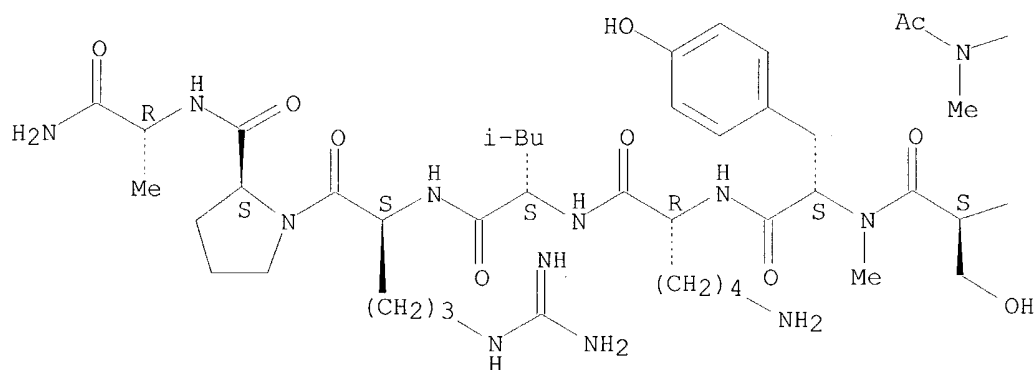
CM 1

CRN 135185-38-5  
 CMF C66 H92 Cl N15 O13

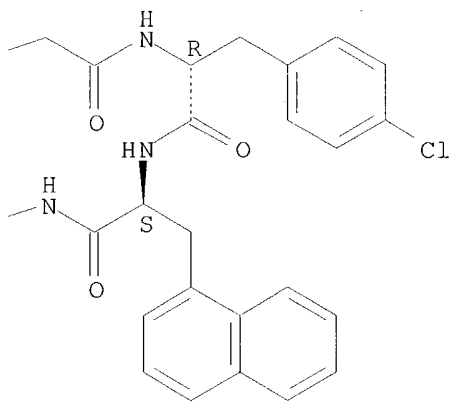
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

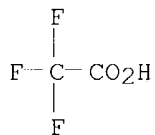


PAGE 1-B



CM 2

CRN 76-05-1  
CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 20 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **135185-37-4** REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-N6-(3-pyridinylcarbonyl)-D-lysyl-L-leucyl-L-arginyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA

INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C72 H95 Cl N16 O14 . x C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

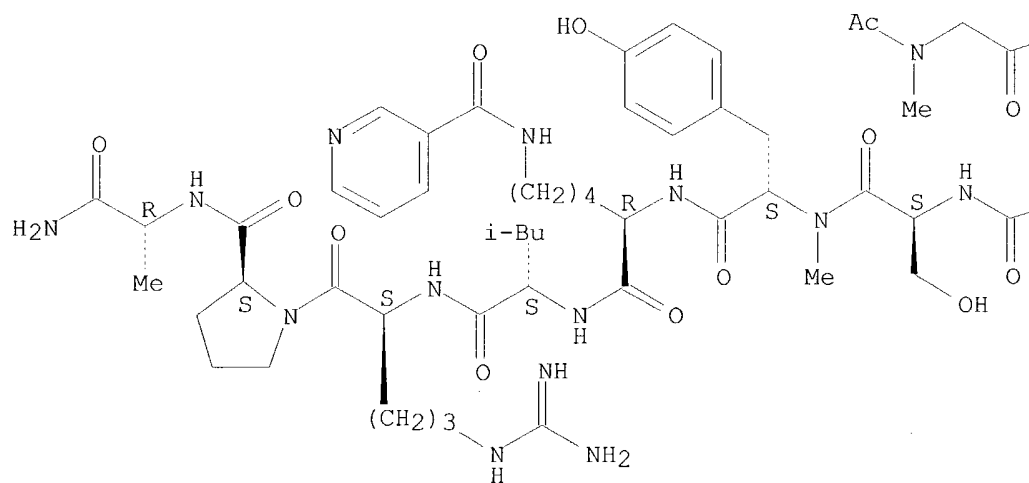
CM 1

CRN 135185-36-3  
 CMF C72 H95 Cl N16 O14

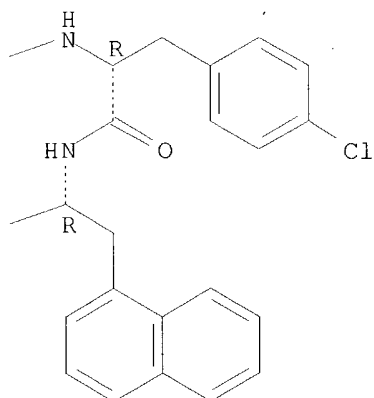
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



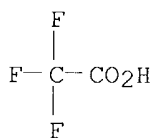
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 21 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **135185-35-2** REGISTRY  
CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-3-(3-pyridinyl)-D-alanyl-L-leucyl-L-arginyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C68 H88 Cl N15 O13 . x C2 H F3 O2  
SR CA  
LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

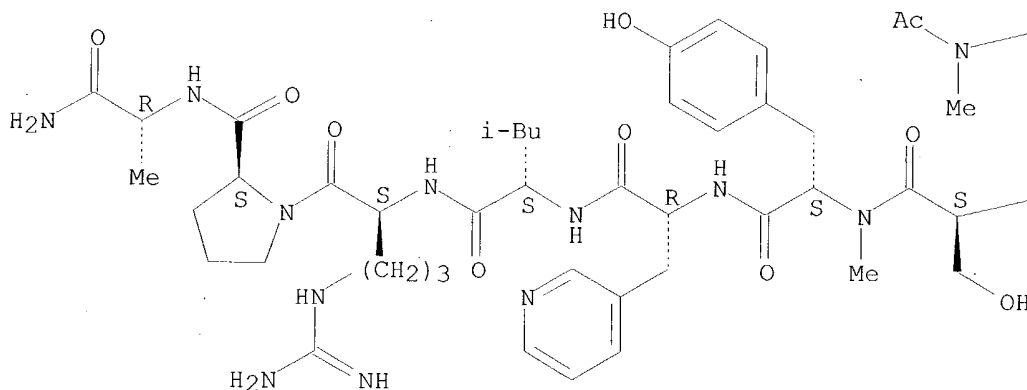
CRN 135185-34-1

CMF C68 H88 Cl N15 O13

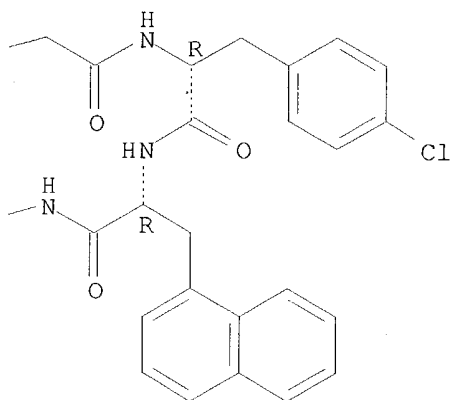
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



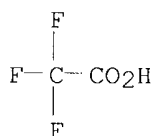
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 22 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 135185-33-0 REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(2-thienyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-D-lysyl-L-leucyl-L-arginyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C60 H88 Cl N15 O13 S . x C2 H F3 O2

SR CA

LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

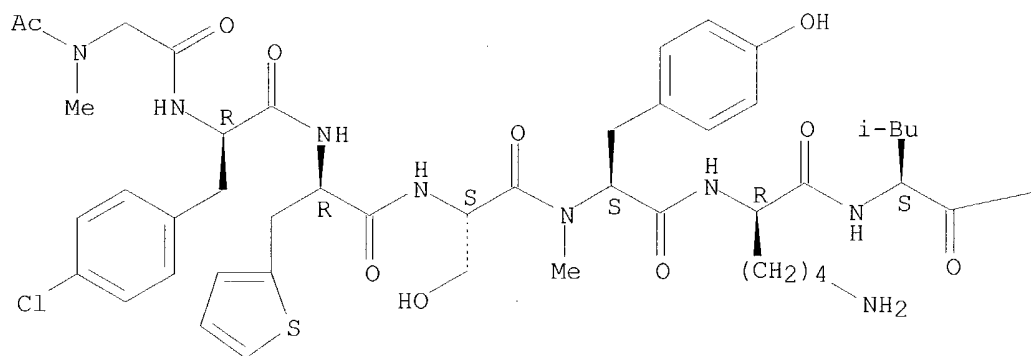
CRN 135185-32-9

CMF C60 H88 Cl N15 O13 S

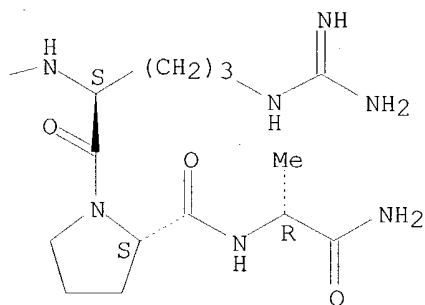
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

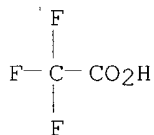


PAGE 1-B



CM 2

CRN 76-05-1  
CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 23 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **135185-31-8** REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-3-(1-naphthalenyl)-D-alanyl-L-seryl-N-methyl-L-tyrosyl-D-lysyl-L-leucyl-L-arginyl-L-prolyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)



FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C66 H92 C1 N15 O13 . x C2 H F3 O2  
 SR CA  
 LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

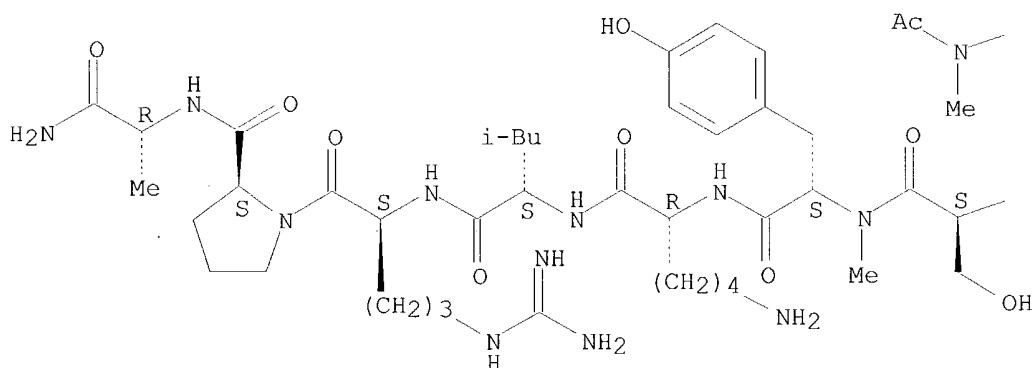
CM 1

CRN 135185-30-7  
 CMF C66 H92 C1 N15 O13

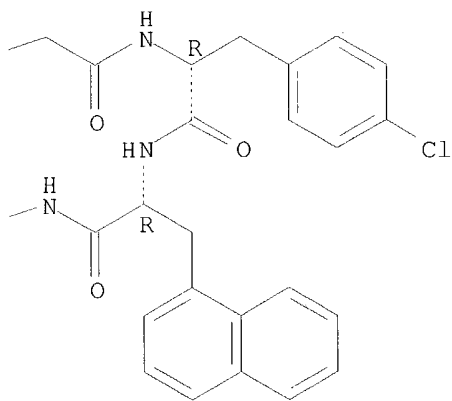
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

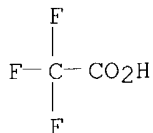


PAGE 1-B



CM 2

CRN 76-05-1  
 CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 24 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **135185-14-7** REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-4-(N-methyl-L-serine)-6-[3-(2-naphthalenyl)-D-alanine]-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-4-(N-methyl-L-serine)-6-[3-(2-naphthalenyl)-D-alanine]-, trifluoroacetate (salt)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C67 H87 N17 O13 . x C2 H F3 O2

SR CA

LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

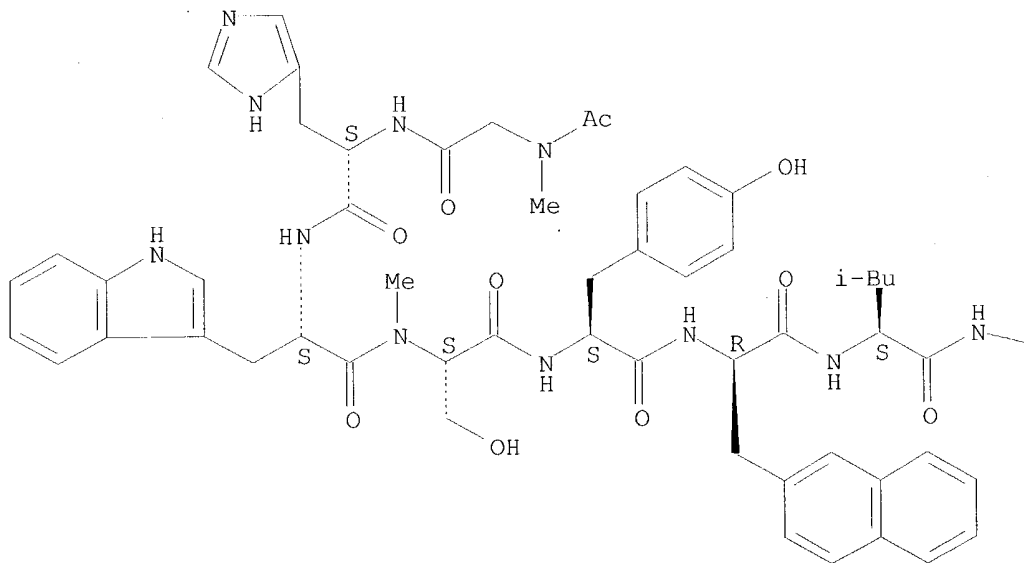
CRN 125323-90-2

CMF C67 H87 N17 O13

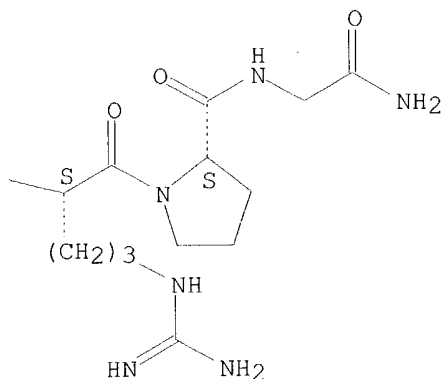
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



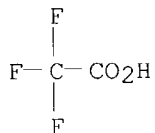
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 25 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **135185-12-5** REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycyl-L-histidyl-L-tryptophyl-N-methyl-L-seryl-L-tyrosyl-D-tryptophyl-L-leucyl-L-arginyl-N-ethyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C65 H87 N17 O12 . x C2 H F3 O2

SR CA

LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

CM 1

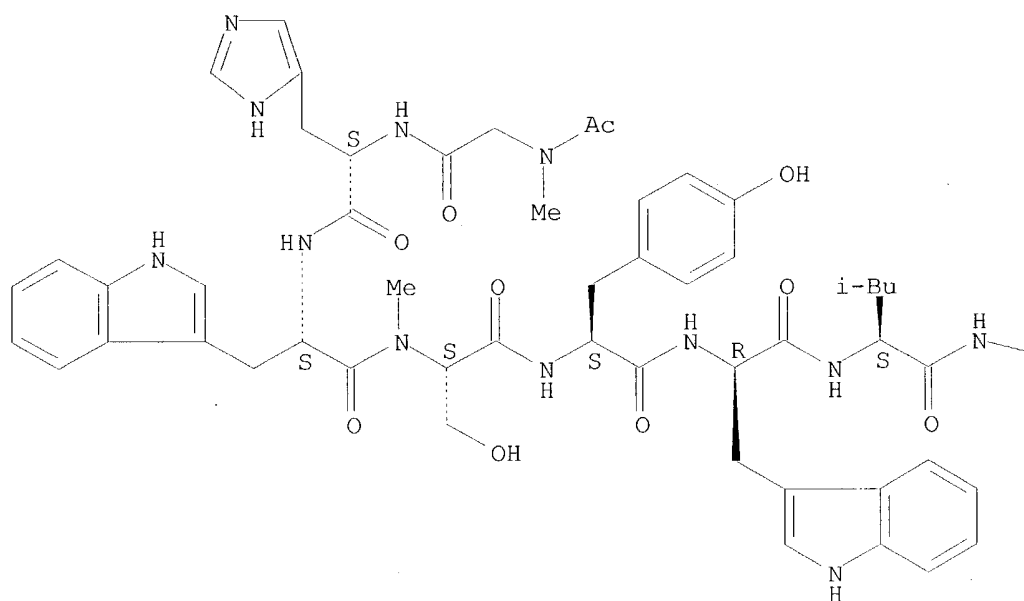
CRN 125323-88-8

CMF C65 H87 N17 O12

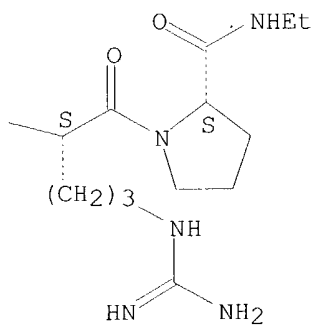
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



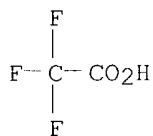
PAGE 1-B



CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 26 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN  
RN **135185-11-4** REGISTRY  
CN L-Prolinamide, N-acetyl-N-methylglycyl-L-histidyl-L-tryptophyl-3-(ethylamino)-N-methyl-L-alanyl-L-tyrosyl-D-tryptophyl-L-leucyl-L-arginyl-N-ethyl-, trifluoroacetate (salt) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE; STEREOSEARCH  
MF C67 H92 N18 O11 . x C2 H F3 O2  
SR CA  
LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

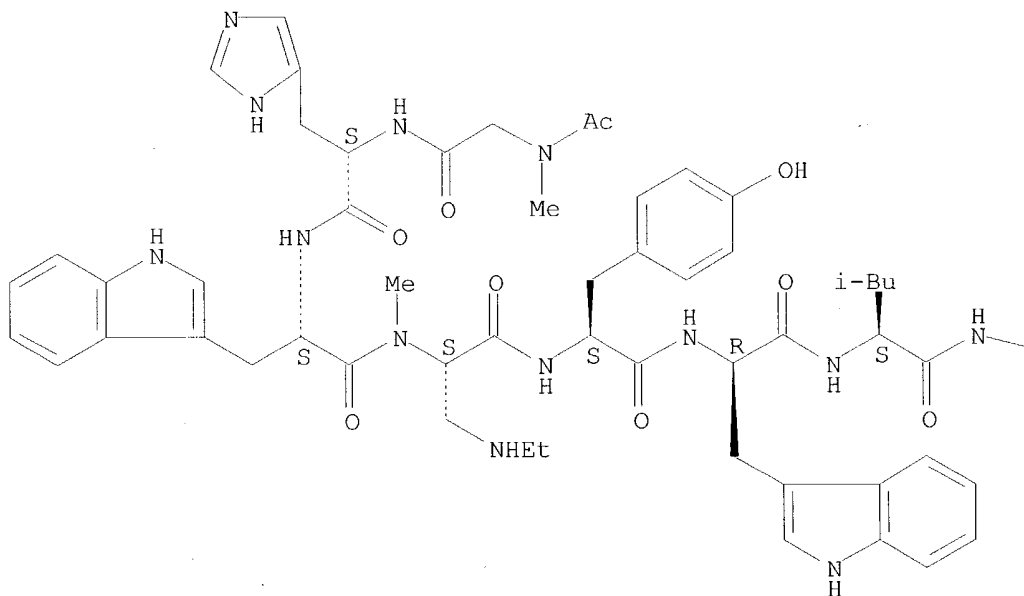
CM 1

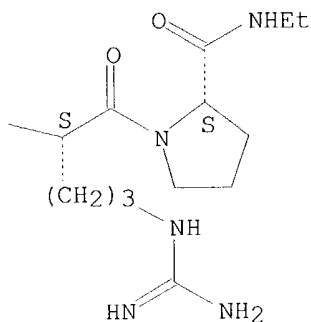
CRN 135185-10-3  
CMF C67 H92 N18 O11

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A

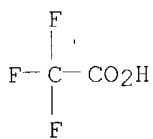




CM 2

CRN 76-05-1

CMF C2 H F3 O2



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

L31 ANSWER 27 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 125323-90-2 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-4-(N-methyl-L-serine)-6-[3-(2-naphthalenyl)-D-alanine]- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-4-(N-methyl-L-serine)-6-[3-(2-naphthalenyl)-D-alanine]-

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C67 H87 N17 O13

CI COM

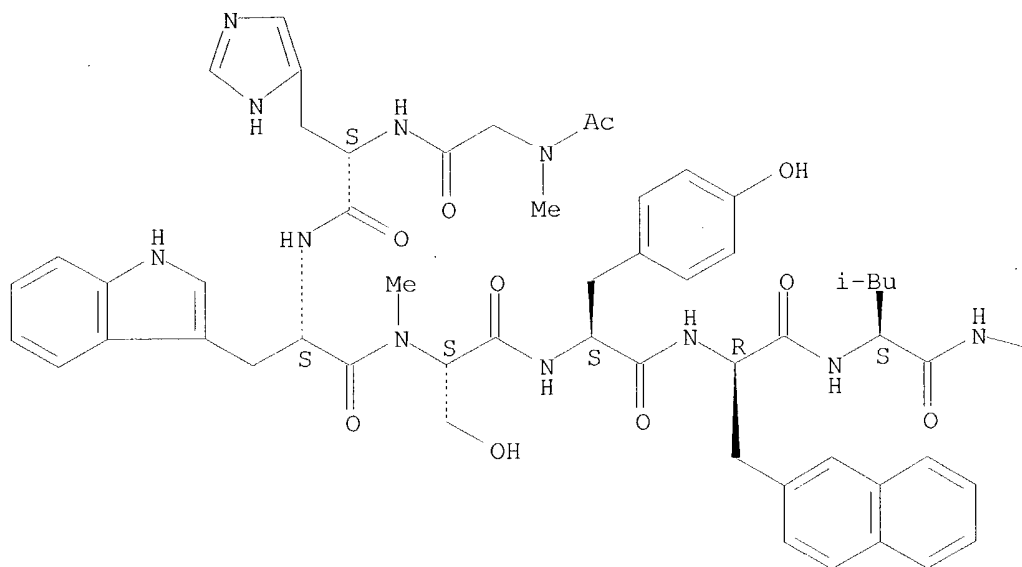
SR CA

LC STN Files: CA, CAPLUS, USPATFULL

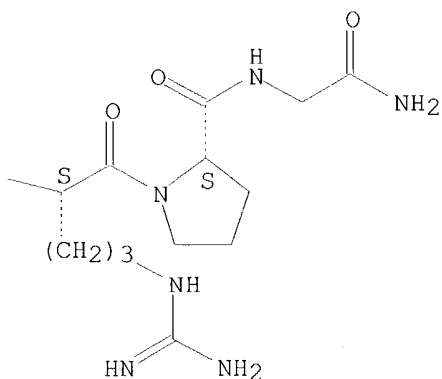
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



2 REFERENCES IN FILE CA (1907 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 115:232885

REFERENCE 2: 112:99259

L31 ANSWER 28 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **125323-88-8** REGISTRY

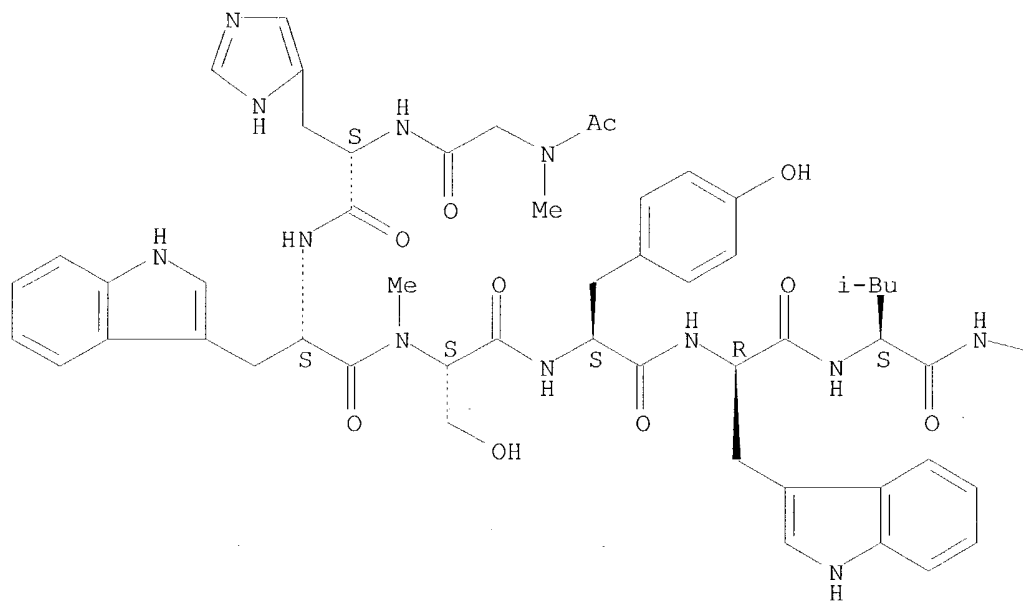
CN L-Prolinamide, N-acetyl-N-methylglycyl-L-histidyl-L-tryptophyl-N-methyl-L-seryl-L-tyrosyl-D-tryptophyl-L-leucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C65 H87 N17 O12  
 CI COM  
 SR CA  
 LC STN Files: CA, CAPLUS, USPATFULL

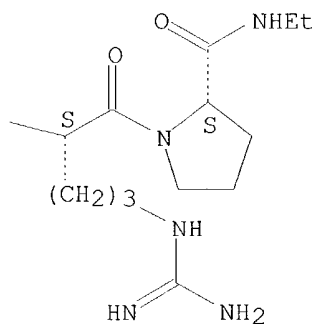
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



3 REFERENCES IN FILE CA (1907 TO DATE)  
 3 REFERENCES IN FILE CAPLUS (1907 TO DATE)



REFERENCE 1: 117:185141

REFERENCE 2: 115:232885

REFERENCE 3: 112:99259

L31 ANSWER 29 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 125323-85-5 REGISTRY

CN L-Prolinamide, N-acetyl-N-methylglycyl-D-phenylalanyl-D-tryptophyl-N-methyl-L-seryl-L-tyrosyl-3-(3-pyridinyl)-D-alanyl-L-leucyl-L-arginyl-N-ethyl- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

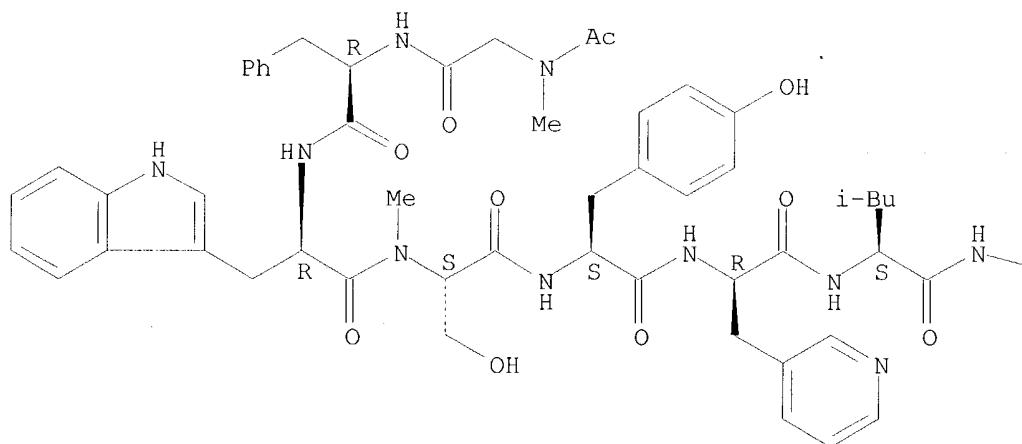
MF C65 H87 N15 O12

SR CA

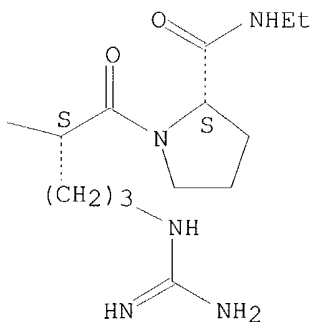
LC STN Files: CA, CAPLUS, USPATFULL

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 112:99259

L31 ANSWER 30 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 102907-98-2 REGISTRY

CN Luteinizing hormone-releasing factor, 1-(N-acetyl-N-methylglycine)-6-D-phenylalanine-9-(N-ethyl-L-prolinamide)-10-deglycinamide- (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C62 H84 N16 O12

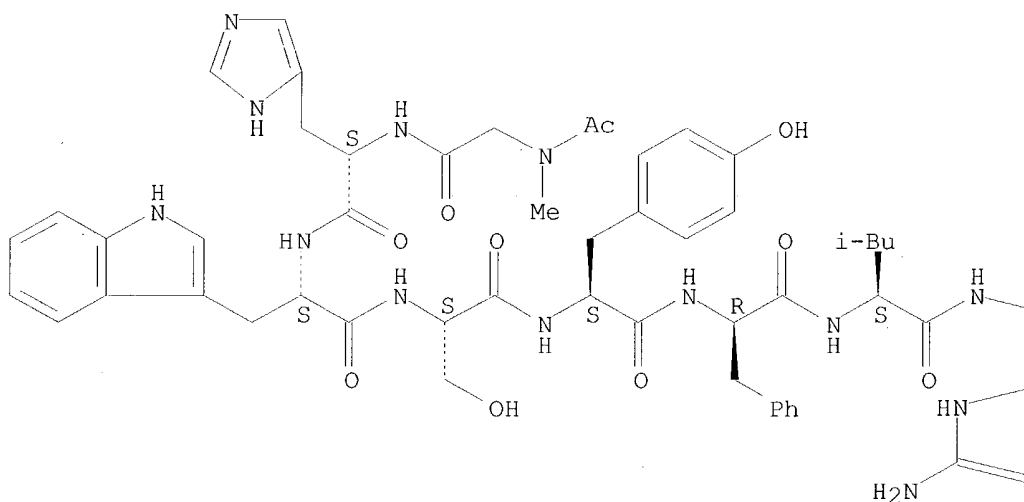
SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

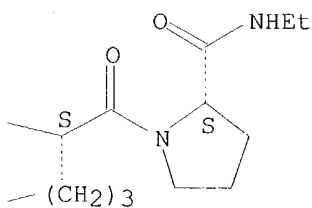
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 105:24631

L31 ANSWER 31 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 102865-91-8 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-tryptophan-9-(N-ethyl-L-prolinamide)-10-deglycinamide- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-tryptophan-9-(N-ethyl-L-prolinamide)-10-deglycinamide-

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C64 H85 N17 O12

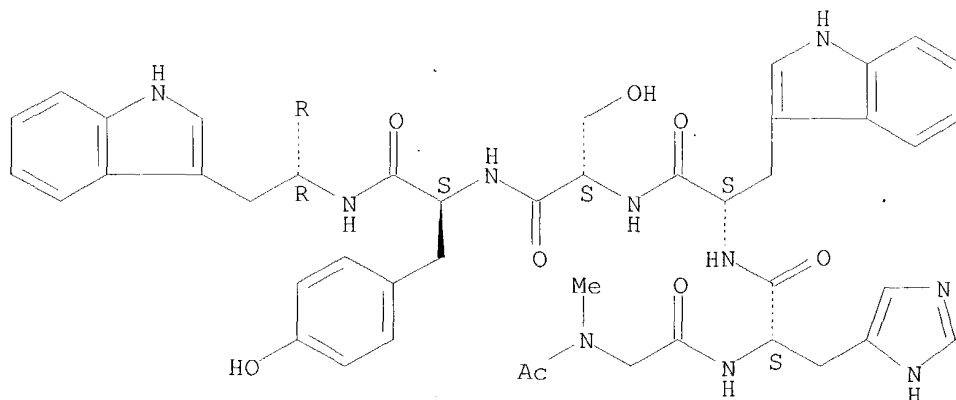
SR CA

LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

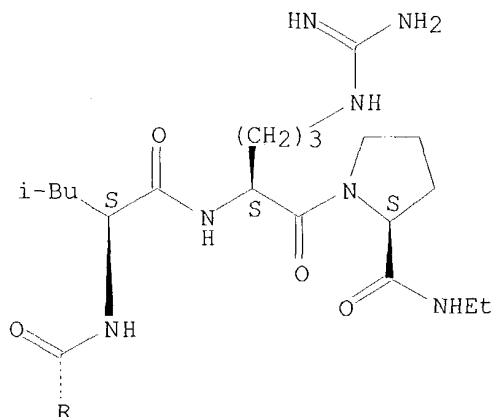
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 105:24631

L31 ANSWER 32 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 102865-90-7 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-tryptophan- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-tryptophan-

FS PROTEIN SEQUENCE; STEREOSEARCH

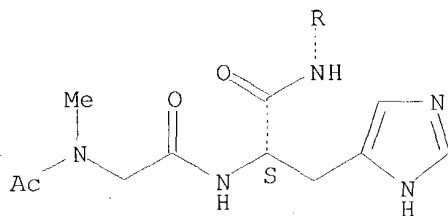
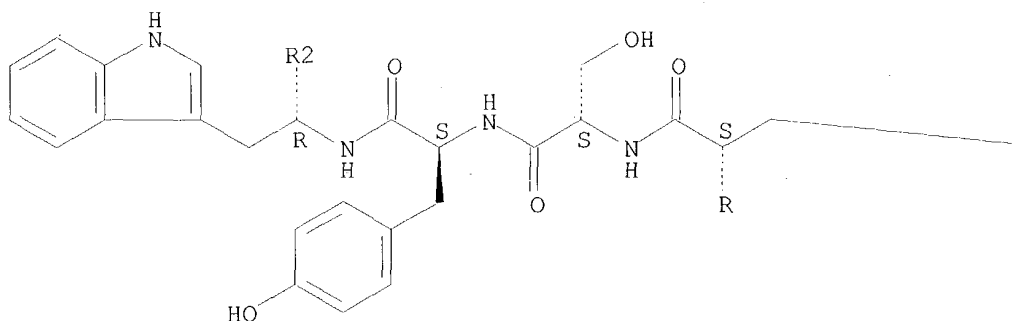
MF C64 H84 N18 O13

SR CA

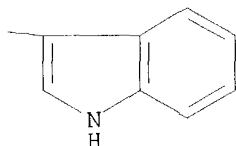
LC STN Files: CA, CAPLUS, TOXCENTER, USPATFULL

Absolute stereochemistry.

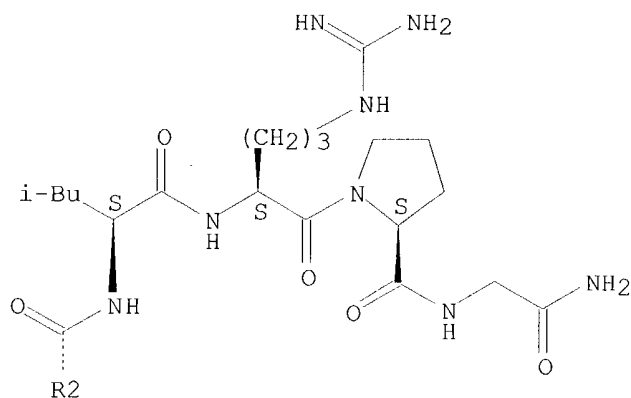
PAGE 1-A



PAGE 1-B



PAGE 2-A



1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 105:24631

L31 ANSWER 33 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **86578-04-3** REGISTRY

CN D-Alaninamide, N-acetyl-N-methylglycyl-4-chloro-D-phenylalanyl-D-tryptophyl-L-seryl-L-tyrosyl-D-phenylalanyl-L-leucyl-L-arginyl-L-prolyl-(9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE; STEREOSEARCH

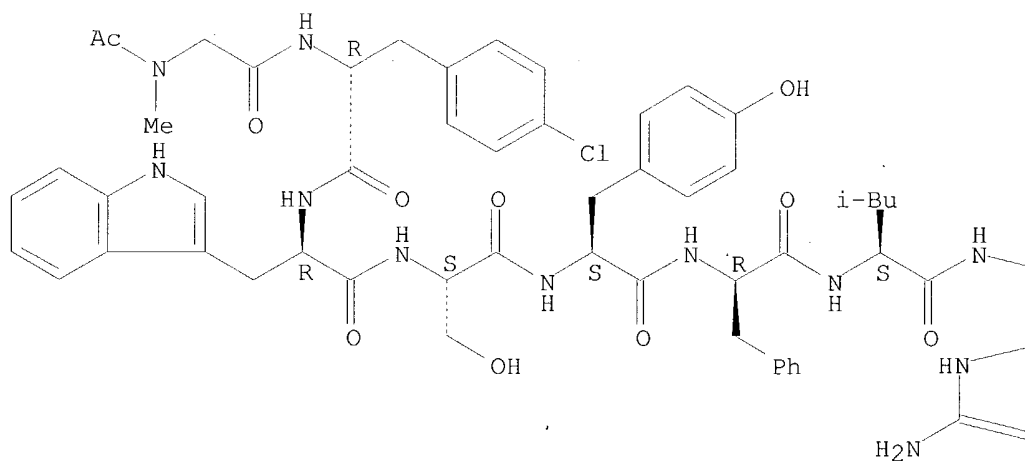
MF C66 H86 Cl N15 O13

LC STN Files: CA, CAPLUS

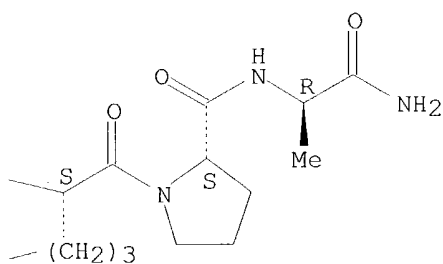
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



2 REFERENCES IN FILE CA (1907 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 101:123209

REFERENCE 2: 99:64462

L31 ANSWER 34 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 81419-14-9 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-  
2-D-phenylalanine-3-D-tryptophan-6-D-tryptophan- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-2-D-phenylalanine-3-D-tryptophan-6-D-tryptophan-

FS PROTEIN SEQUENCE; STEREOSEARCH

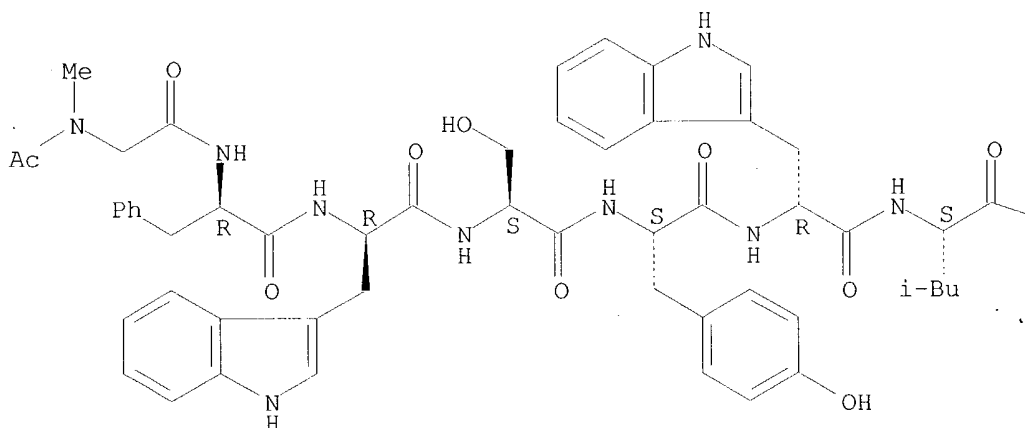
MF C67 H86 N16 O13

LC STN Files: CA, CAPLUS

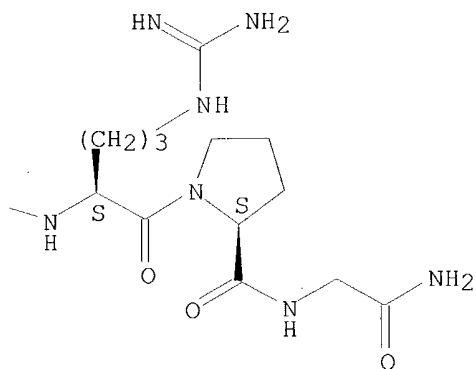
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



1 REFERENCES IN FILE CA (1907 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 96:174580

L31 ANSWER 35 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **62577-30-4** REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-  
 (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-

OTHER NAMES:

CN AcSar1-LH-releasing hormone

FS PROTEIN SEQUENCE; STEREOSEARCH

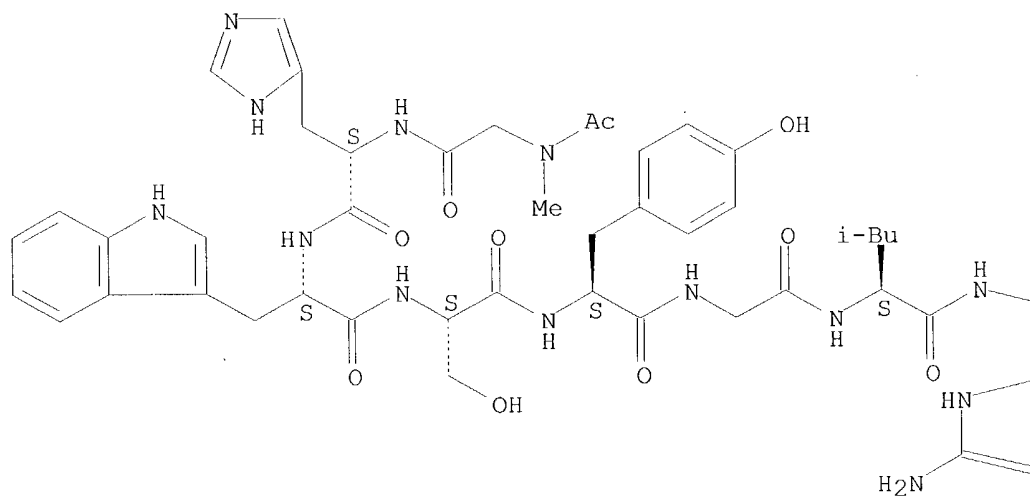
MF C55 H77 N17 O13

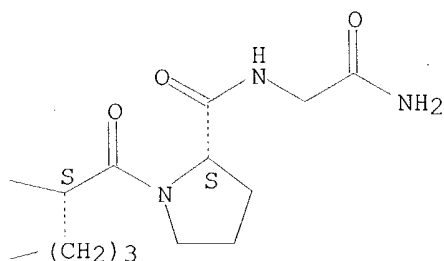
LC STN Files: CA, CAPLUS

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry,

PAGE 1-A





≡NH

2 REFERENCES IN FILE CA (1907 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 89:37043

REFERENCE 2: 86:155951

L31 ANSWER 36 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **58141-55-2** REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-(N-ethyl-L-prolinamide)-10-deglycinamide- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-(N-ethyl-L-prolinamide)-10-deglycinamide-

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C59 H86 N16 O12

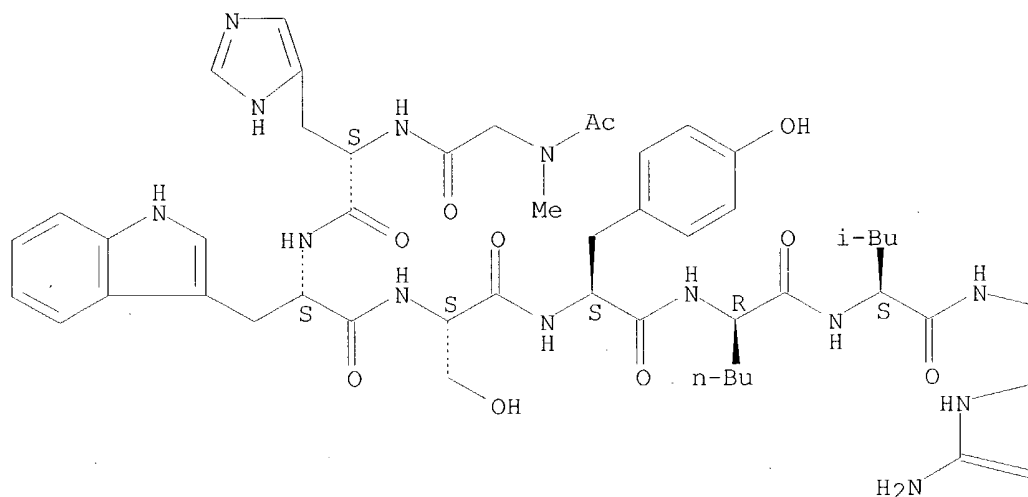
LC STN Files: CA, CAPLUS, IFICDB, IFIPAT, IFIUDB, USPATFULL

**\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\***

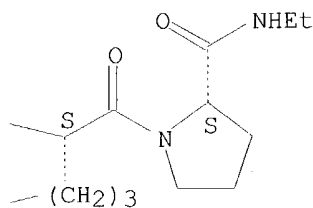
Absolute stereochemistry.



PAGE 1-A



PAGE 1-B



=NH

2 REFERENCES IN FILE CA (1907 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 86:5858

REFERENCE 2: 84:60007

L31 ANSWER 37 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **58083-05-9** REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-[N-(2,2,2-trifluoroethyl)-L-prolinamide]-10-deglycinamide-(9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-[N-(2,2,2-trifluoroethyl)-L-prolinamide]-10-deglycinamide-

FS PROTEIN SEQUENCE; STEREOSEARCH

MF C59 H83 F3 N16 O12

Absolute stereochemistry.

CCCC(=O)N1CCSC1C(=O)NCC(F)(F)F
$$\text{=NH}$$

2 REFERENCES IN FILE CA (1907 TO DATE)  
2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 86:5858

REFERENCE 2: 84:60007

L31 ANSWER 38 OF 40 . REGISTRY COPYRIGHT 2003 ACS on STN

RN 58083-04-8 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-[N-(2-fluoroethyl)-L-prolinamide]-10-deglycinamide- (9CI)  
(CA INDEX NAME)

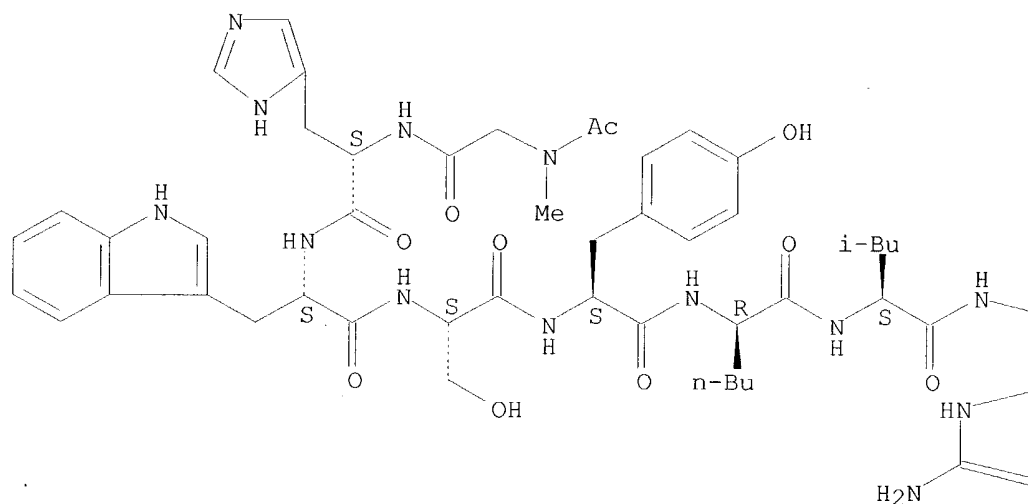
## OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-norleucine-9-[N-(2-fluoroethyl)-L-prolinamide]-10-deglycinamide-  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 MF C59 H85 F N16 O12  
 LC STN Files: CA, CAPLUS, IFICDB, IFIPAT, IFIUDB, USPATFULL

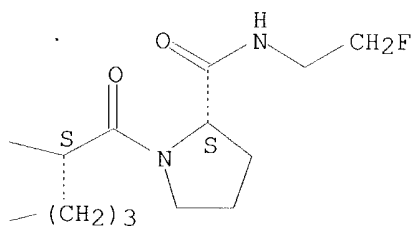
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



=NH

2 REFERENCES IN FILE CA (1907 TO DATE)  
 2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 86:5858

REFERENCE 2: 84:60007

L31 ANSWER 39 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN 58083-03-7 REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-alanine-9-[N-(2,2,2-trifluoroethyl)-L-prolinamide]-10-deglycinamide-(9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-alanine-9-[N-(2,2,2-trifluoroethyl)-L-prolinamide]-10-deglycinamide-

FS PROTEIN SEQUENCE; STEREOSEARCH

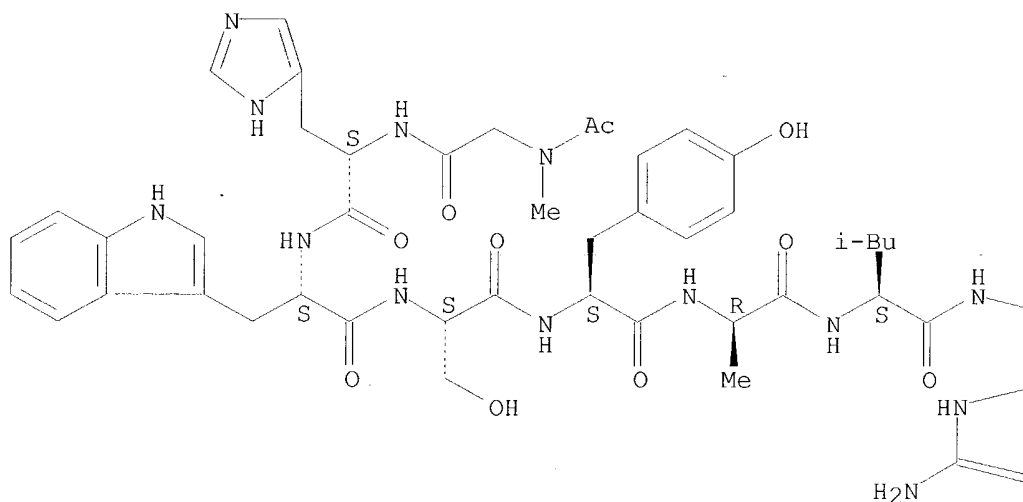
MF C56 H77 F3 N16 O12

LC STN Files: CA, CAPLUS, IFICDB, IFIPAT, IFIUDB, USPATFULL

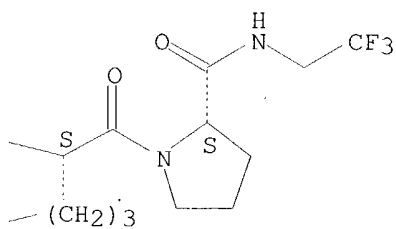
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



=NH

2 REFERENCES IN FILE CA (1907 TO DATE)

2 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 86:5858

REFERENCE 2: 84:60007

L31 ANSWER 40 OF 40 REGISTRY COPYRIGHT 2003 ACS on STN

RN **58083-02-6** REGISTRY

CN Luteinizing hormone-releasing factor (swine), 1-(N-acetyl-N-methylglycine)-6-D-alanine-9-(N-ethyl-L-prolinamide)-10-deglycinamide- (9CI) (CA INDEX NAME)

OTHER CA INDEX NAMES:

CN Luteinizing hormone-releasing factor (pig), 1-(N-acetyl-N-methylglycine)-6-D-alanine-9-(N-ethyl-L-prolinamide)-10-deglycinamide-

FS PROTEIN SEQUENCE; STEREOSEARCH

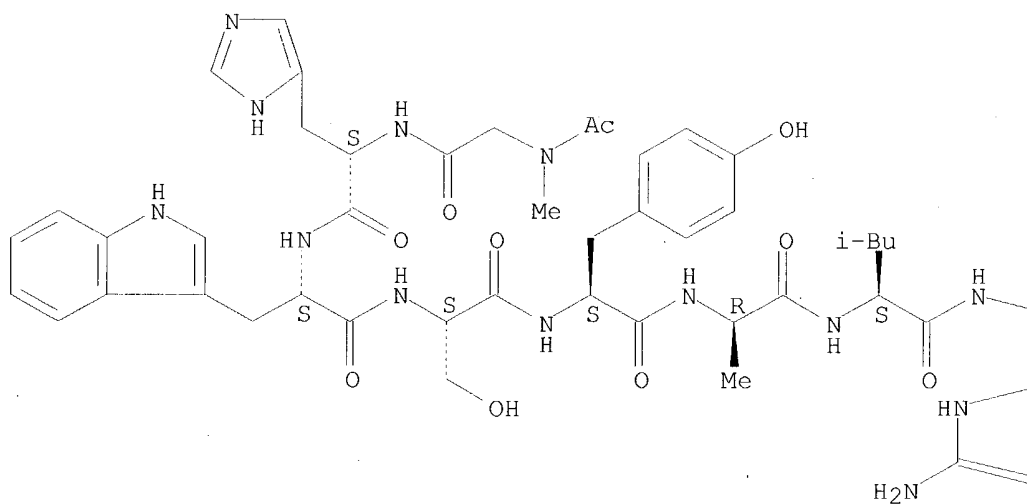
MF C56 H80 N16 O12

LC STN Files: CA, CAPLUS, IFICDB, IFIPAT, IFIUDB, USPATFULL

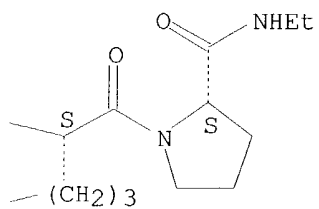
\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

Absolute stereochemistry.

PAGE 1-A



PAGE 1-B



=NH

1 REFERENCES IN FILE CA (1907 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

REFERENCE 1: 84:60007